目　录

第二部分　随机变量

第三部分　随机过程

第一章　事件的概率

§1　概率的概念

当人们谈论某个射手在给定的射击条件下命中率为92%时，这意味着，在一些确定的条件下（同一目标、同一距离、甚至同一步枪等等）他射击了100发子弹，平均约有92发中靶（即意味着近8发脱靶）。当然，也不是每100发都命中92发，有时候命中91发或90发，有时候命中93发或94发，有时命中数甚至可能比92发多得多或少得多，但是，在同一条件下多次重复射击时，这个命中率平均值将仍然不变，一直到发生实质性的变化为止（例如，射手可能提高射击水平，从而使平均命中率达到95%或者更高）。经验表明，该射手在大多数情况下，一百发射击中命中发数接近92。尽管有时在100发射击中，会出现命中数小于88或大于96的情况，但较为罕见。92%这个数字作为射手射击水平的指标通常是不变的，即对该射手而言，在相同条件下，大多数射击命中率几乎是一样的，仅在少数特殊情况下，命中率才明显地偏离平均值。

我们再举一个例子。在某一企业里人们发现，在一定条件下平均有1.6%的产品不符合标准而报废。这意味着，在1000件产品中约有16件不合格。当然，废品有时稍多，有时

稍少，但其平均数接近16，而且在大多数以1000件为一批的各批产品中，废品数也接近于16。不过，这里假定生产条件（如车间工艺流程、设备、原料、工人技能等等）不变。

显然，这类例子不胜枚举。在这些例子中我们看到，在同一类大批量作业（多次射击、大量生产产品等）中，这种或那种形式的重要事件（命中目标、产品不合格等等）的百分比在给定条件下几乎总是相同的，仅仅在少数情况下才大大地偏离某一平均数值。因此可以说，这个平均值是该大批量作业在严格规定的条件下的特征值。命中百分比用来描述射手的射击技术，废品百分比用来评价产品质量优劣程度。由此可见，这个数值对于军事、技术、经济、物理、化学等等各个方面有十分重要的意义。它不仅能使我们评价已经出现的大量的现象，而且可以预测这种或那种大批量作业的未来结局。

如果射手在给定射击条件下，100发中平均 92发 命中目标，则我们说该射手 在 给 定 射击条件下的命中概率为92%（或92/100，或0.92）。如果在给定条件下，某一企业每1000件产品中平均有16件报废，则说该产品出现废品的概率为0.016或1.6%。

在某个大批量作业里，究竟把什么称为事件的概率呢？这个问题现在不难回答。大批量作业总是大量重复彼此类似的单个作业（射击——多次重复单个的发射；批生产——大量生产单个的产品等等）。我们对单个作业的结果（单个射击的命中情况，单个制品的不合格情况等等）很感兴趣，而且首先对这种或那种大批量作业中这些结果的数值（多少发命中目标，多少产品将报废）感兴趣，于是，把某个大批量作业

中这类感兴趣的、“期望的”㊀结果的百分比(或比率)称为该结果的**概率**。这里应该经常注意，这个或者那个事件（结果）的概率仅在准确给定的条件下进行大批量作业才有意义。这些条件的一切实质性变化通常都会引起概率的变化。

如果有这样的一个大批量作业，事件A(例如命中目标)在b次单个作业(射击)中出现了a次，则事件A在给定条件下的概率是a/b(或$\frac{100a}{b}$%)。因此，可以把这些“**期望的**”**结果数**与组成大批量作业的**所有单个作业数**的平均**比值**称为单个作业“期望的”**概率**。不言而喻，如果某个事件的概率等于a/b，那么，在b次单个作业中，这个事件出现的次数可能大于a，也可能小于a，仅仅是该事件发生的平均次数约为a；而在大多数情况下，在b次单个作业中，**特别是当b是个很大的数时**，事件A出现的次数将接近于a。

例1 某城市在第一季度里人口出生情况是：

一月份—145个男孩和135个女孩；

二月份—142个男孩和136个女孩；

三月份—152个男孩和140个女孩。

男孩出生的概率有多大呢？男孩出生率是：

一月份 $145/280 \approx 0.518 = 51.8\%$；

二月份 $142/278 \approx 0.511 = 51.1\%$；

三月份 $152/292 \approx 0.521 = 52.1\%$。

我们看到，在三个月内男孩出生的算术平均值接近于$0.516 = 51.6\%$，在该条件下所求的概率约为0.516或

㊀ 在第二个例子里，似应称为“不期望的”。不过，在概率论中，直接把那些能出现人们解题时感兴趣事件的结果称为“期望的”。

51.6%。这个数字在研究人口动态情况的科学——人口学里已为人们所熟知了。它表明，在通常条件下，就是在不同时期里，男孩的出生率不会过分偏离这个数值。

例2　上个世纪初发现了著名的布朗运动（以发现者、英国植物学家布朗姓名命名）自然现象。这个现象是悬浮在流体中㊀的物质微粒在没有明显外因影响的情况下作不规则运动。

当气体动力理论没给出简明而准确的解释（即悬浮在流体中的粒子运动是这些粒子与流体分子碰撞的结果）之前，显然人们无法弄清楚这种自发的不规则运动的原因。气体动力理论能使人们统计计算在某一定量流体中，或者无一个，或者有一个、两个、三个……悬浮粒子的概率。为检验理论结果。人们曾经做过一系列试验。

我们援引瑞典物理学家斯费德别尔格对悬浮在水中的金微粒观察518次的结果。发现在被观察的部分空间里，未见到一个微粒有112次，见到1个微粒有168次，见到2个微粒有130次，见到3个微粒有69次，，见到4个微粒有32次，5个——5次，6个——1次，7个——1次，

由此得出：

发现0个微粒的次数比率是112/518＝0.216；

发现1个微粒的次数比率是168/518＝0.324；

发现2个微粒的次数比率是130/518＝0.251；

发现3个微粒的次数比率是69/518＝0.133；

㊀　即处于随遇平衡状态。

发现 4 个微粒的次数比率是32/518＝0.062；
发现 5 个微粒的次数比率是 5/518＝0.010；
发现 6 个微粒的次数比率是 1/518＝0.002；
发现 7 个微粒的次数比率是 1/518＝0.002；
由此可见，观察结果与理论预测概率恰好吻合。

例3 在许多重要的实际问题里，需要知道哪些俄文字母用得较多。例如，在配置印刷铅字盒时，所有字母都按同一数量储备是不合理的，因为某些字母比其他一些字母在文章里出现的次数要多得多。为此，必须使经常出现的字母数量多一些。对文学作品字母研究的结果可使人们估计出俄文字母（包括词间空格）出现的频率，其频率表㊀如下（按相对频率依次减小的顺序排列）。

研究表明，在任意选取的一段包含一千个字母与空格的文章中，字母Ф出现 2 次，字母K——28次，字母O——90次，而空格出现175次。这些数据对于配置印刷铅字盒是相当宝贵的。

近些年来，这类研究工作已经不只局限于统计俄语文章中的字母，而且开始广泛地用于阐明俄语的特点，乃至不同作者的文学风格了。此外，可用电报通讯数据编制出比较经济的电报代码，便于用少量符号及较快速度传递信息。须知，目前所采用的电报代码还不是最经济的。

㊀ 此表引自А·М·Яглом和И·М·Яглом的优秀科普读物《概率与信息》 莫斯科科学出版社，1973年。

字　母	空格	О	е,ё	А	И	Т	Н	С
相对频率	0.175	0.090	0.072	0.062	0.062	0.053	0.053	0.045
字　母	Р	В	Л	К	М	Д	П	У
相对频率	0.040	0.038	0.035	0.023	0.026	0.025	0.023	0.021
字　母	Я	Ы	З	Ь,Ъ	Б	Т	Ч	Й
相对频率	0.018	0.016	0.016	0.014	0.014	0.013	0.012	0.010
字　母	Х	Ж	Ю	Ш	Ц	Щ	Э	Ф
相对频率	0.009	0.007	0.006	0.006	0.004	0.003	0.002	0.002

§2 不可能事件与必然事件

显然，事件的概率总是正数或零。由于在表示概率的分数中分子不可能大于分母（即“期望的”作业数不可能大于全部作业数），因此概率不可能大于1。假定用 P(A) 表示事件A的概率，不管该事件是什么样的，其概率为

$$0 \leqslant P(A) \leqslant 1$$

P(A)越大，事件A出现的次数越多。例如，射手命中目标的概率越大，说明他射击成功的机会越多，射击水平越高。如果事件的概率很小，则事件A很少出现，如果P(A)=0，则事件A或者不出现，或者出现的次数极少，以致实际上可认为这个事件不可能出现。相反，如果P(A)接近于1，那么，在表示事件A的概率的分数中，分子数接近于分母数，说明这种事件在大多数情况下出现，或者说绝大多数作业都是“期望的”。如果P(A)=1，则事件A“总是”或“几乎总是”出现。因此，事实上可认为事件A是**必然事件**，即准确地认定要出现的事件。如果P(A)=1/2，则事件A出现的次数约占全部情况的一半，这意味着可把“期望的”作业与“不期望的”作业近似地看作一样多。如果P(A)>1/2，则事件A出现的次数比不出现的次数要多，当P(A)<1/2时，情况正好相反。

事件的概率究竟小到多少才可认为它是不可能事件呢？回答这个问题不能一概而论。这完全取决于事件的重要程度。比如，0.01是一个不大的数。如果有一批炮弹，而0.01

是炮弹落下未爆炸的概率。这就意味着约有1%的射击无效。这是可以容许的。但是，如果0.01是跳伞时未开伞的概率，当然这无论如何是不能容许的。这些例子表明，在每个具体问题里需要根据实际情况预先确定：某个可不予考虑的事件的概率应为多小才不致损害事业。

§3 题 解

问题 一个射手命中率为80%，另一个射手（在同一射击条件下）为70%，如果两个射手同时射击目标，求射中目标的概率。两人同时射击时，两发子弹尽管只有一发命中目标，也认为目标被射中。

第一种解法。设进行100次双人射击。其中，第一个射手80次射中目标，其余20次脱靶。由于第二个射手100次中平均有70次射中目标，即10次中有7次射中目标，则可以期望，在第一个射手射击脱靶的那20次中，第二个射手成功射击目标约有14次。这样一来，在全部一百次射击中，目标被射中的次数约为80＋14＝94次。因此，两个射手同时射击目标时，射中目标的概率等于94%或0.94。

第二种解法。假设仍进行100次双人射击，并且已经知道，在射击中第一个射手约有20次脱靶。由于第二个射手在100次射击中约有30次脱靶，即在10次中有3次脱靶，则可以期望，在第一个射手脱靶的20次射击中，第二个射手射击时也会有6次脱靶。在这6次中每次射击都不命中目标，在其余94次中，每次至少有一个射手射中目标，即目标被射中。由此再次得出结论，在双人同时射击时，100次中有94次机会射

中目标，即命中概率为94%或0.94。

上述问题虽很简单，但它使我们得出了一个很重要的结论：往往有这样一些情况，根据一些事件的概率就能求出另一些较复杂事件的概率。事实上，这类情况很普遍，它不仅在军事活动中存在，而且在所有科学及所有涉及大批量现象的实践活动中存在。

当然，对这一类中每个新问题都要求出特殊的解是很不方便的。科学总是力求建立一般规则，利用这些规则就可以机械地或近乎机械地求解其它彼此类似的问题。

研究大批量现象的一般规则的学问叫做**概率论**。本书将阐述这门学科的一些基本原理。

概率论象算术或几何一样是数学中的一个分支。因此，它的研究方法是精确地推理，使用的工具是公式、表格、图形等等。

第二章 概率的加法定理

§4 概率的加法定理导出

加法定理是概率计算中最简单、最重要的一条常用基本定理。

在射击图1所示的靶标时，每个射手在一定距离上击中靶标上1、2、3、4、5、6区间的概率各不相同。假定某个射手击中1区间的概率是0.24，击中2区间的概率是.17。这就意味着该射手平均射击的100发子弹中，有24发子弹射中1区间，有17发子弹射中2区间。

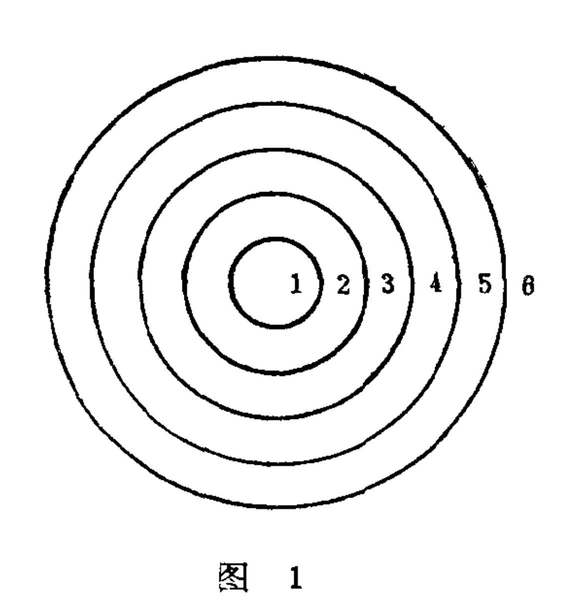


图 1

假设在某次比赛中，子弹落在1区间为射击优秀，落在2区间为射击良好，那么，射手的射击成绩为优秀或良好的

概率是多大呢？

这个问题不难回答。射手每射击100发子弹，约有 24 发落在 1 区间，17发落在 2 区间。在每100发子弹中显然有24 + 17 = 41发子弹或者落在 1 区间，或者落在 2 区间。因此，所求概率等于0.24 + 0.17 = 0.41。于是，射手成绩为优秀或良好的概率等于优秀的与良好的射击概率之和。

再看另一个例子。乘客在某电车站等待26路或 16 路 电车，该站停靠16、22、26、31四路电车。假定各路电车停靠的频率一样，求乘客期待的电车首先停靠的概率。

显然，16路电车首先停靠的概率等于1/4，26路车 也 是1/4，于是所求概率等于1/2，即

$$1/4 + 1/4 = 1/2$$

因此可以说，16路电车或26路电车首先停靠的概率等于16路电车和26路电车首先停靠的概率之和。

现在可以进行一般性描述。

在进行某项大批量作业时，已知，在由b个单独作业组成的每批作业里平均出现以下现象：

某种事件A_1出现a_1次

某种事件A_2出现a_2次

某种事件A_3出现a_3次

…………………………

换句话说，

事件A_1的概率等于a_1/b

事件A_2的概率等于a_2/b

事件A_3的概率等于a_3/b

…………………………

在某个单独作业里出现A_1、A_2、A_3……事件中任意一种的概率有多大呢？我们所关心的事件可能是A_1，或A_2，或A_3…，它在由b个单独作业组成的批量作业里出现$a_1+a_2+a_3+\cdots$次，则所求的概率为

$$a_1/b+a_2/b+a_3/b+\cdots=\frac{a_1+a_2+a_3+\cdots}{b}$$

这可写成如下公式：

$$P(A_1，或A_2，或A_3\cdots)$$
$$=P(A_1)+P(A_2)+P(A_3)+\cdots$$

无论是在上述实例中，还是在一般性推论中，总是假定：所研究的任意两个（例如A_1和A_2）结果彼此是不相容的，即它们在同一单个作业中不可能都出现。到站的电车不可能既是某乘客期待的同时又不是他所期待的线路车。它或者满足乘客的愿望，或者不满足。

我们关心的各个结果彼此不相容的假设十分重要。无此假设，加法定理就不正确；而用它计算势必酿成大错。例如，第一章第三节中的问题，该题恰恰是求在双人射击时，或者第一个射手命中目标，或者第二个射手命中目标的概率。而且第一个射手的命中概率为0.8，第二个射手的命中概率为0.7。假如用加法定理来解这一问题，立即可以算出所求概率等于0.8+0.7=1.5。这个结果显然是荒谬的，因为事件的概率不能大于1。得出这个错误结果的原因是，在不该采用加法定理的地方采用了加法定理。这个问题中涉及的两种结果是彼此相容的，因为两个射手在同一次双人射击时完全可能有都命中目标的情况，开始计算概率时所犯的主要错误就在于错误地应用了加法定理；因此在每次使用加法

定理时，必须提防这种错误。在拟采用加法定理的诸事件中，请检查一下它们彼此之间是否真是互不相容。

现在可以给出加法定理的一般性表述。

加法定理　在某一批量作业中，如果其结果A_1，A_2，……A_n俩俩互不相容，则出现其中任意一种结果（无论是哪一种结果）的概率等于每个结果的概率之和。

§5　全事件系统

在第三次发放公债的二十年有效期里，有三分之一的债券中奖，而其余三分之二的债券无奖，只按债券票面价值还本。换言之，就这次公债而论，每张债券的中奖概率是1/3，不中奖的概率是2/3。中奖与不中奖是**对立事件**，即就每张债券而言，中奖与不中奖这两个对立事件只有一个且仅有一个事件必然出现。它们的概率之和为：

$$1/3+2/3=1$$

这是必然的。一般说来，如果A_1和A_2是两个对立的事件，且如果在b次作业中事件A_1出现a_1次，事件A_2出现a_2次，显然$a_1+a_2=b$，则有

$$P(A_1)=a_1/b,\ P(A_2)=a_2/b$$

于是

$$P(A_1)+P(A_2)=a_1/b+a_2/b=\frac{a_1+a_2}{b}=1$$

从加法定理也可得到同样的结果：因为对立的事件互不相容，则有

$$P(A_1)+P(A_2)=P(A_1\text{或}A_2)$$

从对立事件的定义可知，事件（A_1或A_2）必定出现，因此事件（A_1或A_2）是必然事件，其概率等于1，于是我们重又得到

$$P(A_1)+P(A_2)=1$$

两个对立事件的概率之和等于1。

这个规则可以导出一个用同一方法能够证明的一个重要通则。设有n个（任意数）事件A_1、A_2、…A_n，在每一单独作业中，这些事件中有一个且仅有一个事件必定出现，我们把这组事件称为**全事件系统**。显然，全事件系统是由所有对立事件组成。

组成全事件系统的各事件的概率之和等于1。

事实上根据全事件系统的定义，该系统诸事件中任意两个事件是互不相容的，所以加法定理为

$$P(A_1)+P(A_2)+\cdots+P(A_n)=P(A_1, 或A_2,\cdots,或A_n)$$

不过，这个等式右边部分是必然事件的概率，其概率值为1，因此，就全事件系统而言，

$$P(A_1)+P(A_2)+\cdots+P(A_n)=1$$

由此证得。

例1 射手射击图1所示的靶标，每100发射击成绩为：

44发命中1区，

30发命中2区，

15发命中3区，

6发命中4区，

4发命中5区，

1发命中6区，

（44+30+15+6+4+1=100）。显然，这六种射击结果组成了全事件系统。各部分对应的概率等于

0.44　0.30　0.15　0.06　0.04　0.01

我们可得到

$$0.44+0.30+0.15+0.06+0.04+0.01=1$$

有时，射到第6区间的子弹全部或部分未打在靶子上，从而无法统计。但是这不妨碍求命中该区间的概率，在这种情况下只要用1减去命中其它各区间所有概率之和就是了。

例2 统计表明，在某一纺织工厂中，纺织机每一百次停车中需要纺织工人动手操作的情况平均是：

22次由于断了经纱，

31次由于断了纬纱，

27次由于换梭子，

3次由于打梭板损坏，

而其他停车是由于别的缘故。

我们发现，除了其他原因以外，纺织机停车有四种确实的原因，其概率分别等于

0.22　0.31　0.27　0.03

这些概率之和等于0.83。上述纺织机停车的原因与其他原因一起组成了全事件系统。因此，其他原因引起纺织机停车的概率为$1-0.83=0.17$。

§6　实　例

根据上述全事件系统的理论，时常有效地进行所谓“先验的”概率计算。例如，我们来研究一下宇宙质点落到一小

块被等分为六个方格的矩形场地上的情况（图 2）。这些场地处于相同的条件之下，没有任何根据作出宇宙质点落到六个方格中某一格的机会比落到其他方格的要多一些的假定。因此假定：宇宙质点落到六个方格中每一格的机会平均一样多，即质点落到这些方格的概率p_1、p_2、p_3、p_4、p_5、p_6彼此相等。假如仅只关心落到该场地的宇宙质点，则由此得出，每一格的概率P等于1/6，由于这些概率的数值彼此相等，因此根据前面业已证明的理论，概率之和等于1。

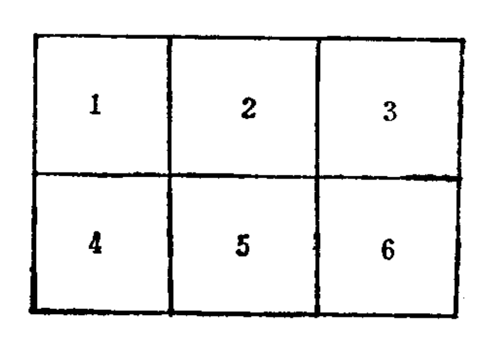


图 2

当然，根据一系列假设所得出的结论需要实际经验的检验。但是，在同样的情况下，我们是如此习惯于肯定经验结论，以致在实际检验之前完全有把握地相信我们的理论假设。在这样一些情况下，通常说该问题可以有n个不同的等概率结果（如在本例中，可以把宇宙质点落到六个方格中之一格的结果作为宇宙质点落到矩形场地的结果）。

n个结果中，每一独立结果的概率等于1/n。这种“先验”计算的重要性在于，在很多情况下，即在大批量作业的概率或者完全不可能确定，或者特别难于确定的条件下，它可以预测出事件的概率。

例1 公债券的顺序号通常用五位数字表示。求随意抽取的中奖序号末位数字为7（例如，序号59607）的概率。为此，根据概率的定义，必须查看长长一系列中奖号码表并统计出有多少中奖序号末位数字为7。该数字与中奖总数之比就是所求的概率。但是，0、1、2、3、4、5、6、7、8、9十个数字中每一个数字毫无例外地有同样多的机会作为中奖序号的末位数字。因此毫不犹豫地推测出所求的概率等于0.1。这个理论“预测”的正确性极易检验：对任意一段中奖号码表作一些必要的统计就可证实，10个数中任一个数出现在末位的机会确实占总数的1/10。

例2 连接相距2公里的A、B两站的电话线在某一不明地点断线了。求电线在离A站450米内断线的概率？设想把全部电线按一米分段，由于所有这些线段完全一样，就可以假定每米电线断线的概率是一样的，由此很容易地算出，所求概率等于

$$450/2000=0.225$$

第三章　条件概率与乘法定理

§7　条件概率的概念

假设有两个工厂生产电灯泡，第一个工厂提供全部所需产品的70%，第二个工厂提供30%。第一个工厂平均每一百个电灯泡中有83个符合标准㊀，而第二个工厂平均每一百个电灯泡中只有63个符合标准。

根据这些数据极易计算出，被用户购买的平均每一百个电灯泡中，有77个符合标准㊁。因此，购买符合标准的电灯的概率等于0.77。不过，假定已经查明商店里所有的电灯泡都是第一个工厂生产的。那么用户购买符合标准的电灯泡的概率就要改变。它将等于

$$83/100=0.83$$

这个例子表明：给某个作业（本例中是指购买电灯泡）赖以发生的总条件里补充一个新条件（本例中的新条件是指明电灯泡是由哪个工厂生产的），就可能改变单个作业的这种或那种结果。显然，根据概率定义本身要求，进行大批量作业的总条件应准确确定。如果把某个新的条件补充到总条件之中，实质上就改变了这个总条件。于是，大批量作业就是在新条件下进行。这实际上已成为另一个作业了。在这种

㊀　如果电灯泡能亮1200小时以上，则称它符合标准（满足标准要求）；反之，则称它不符合标准。

㊁　即是$0.7\times83+0.3\times63=77$

情况下，这种或那种结果的概率已不是在初始条件下的概率了。

这样一来，对于购买电灯泡这同一事件有两个不同的概率，不过这两个概率是在不同条件下计算的：未增加补充条件（不考虑电灯泡是哪个工厂生产的）时，得到购买符合标准的电灯泡的**无条件概率**为0.77；增加补充条件（即电灯泡是第一个工厂生产的）时，得到与前面结果有所不同的**条件概率**为0.83。如果用事件A表示购买符合标准的电灯泡，用事件B表示电灯泡是第一个工厂生产的，则通常用$P(A)$表示事件A的无条件概率，用$P_B(A)$表示事件A在事件B存在的条件下（即电灯泡是第一个工厂生产的）的概率，从而得到

$$P(A)=0.77 \qquad P_B(A)=0.83$$

由于只有在某些准确确定的条件下才能论及给定作业的某种结果的概率，所以严格地说，一切概率都是条件概率。因此无条件概率这个词本身不可能存在。不过大多数具体问题的情况是：某个确定的、完成所有作业都需要的总条件K是该问题中一切作业的基础条件。如果在计算某概率时，除了总条件K以外，不附加任何其他条件，那么可以把这种概率称为**无条件概率**；如果计算概率时除了所有作业都需要的总条件K外，还有某些准确确定的补充条件，那么这种概率就称为**条件概率**。

当然在上例中假定，出售的全部电灯泡都是在同样的确定条件下生产的。这个假定理所当然，以致在提出问题时都无需赘述。如果对该种电灯泡不附加任何补充条件，那么把电灯泡试验的这种或那种结果的概率称为无条件概率。如果再

附加某种补充要求，那么在这些条件下所计算的概率就是条件概率了。

例1 在本节开头描述的问题里，销售第二个工厂的电灯泡的概率显然等于0.3。现假定销售的电灯泡都符合质量标准，那么销售第二个工厂生产的符合标准的电灯泡的概率怎样呢？

销售的每一千个电灯泡里有770个质量符合标准，其中第一个工厂的电灯泡581个，第二个工厂的电灯泡189个㊀。因此，销售第二个工厂生产的符合标准的概率是

$$189/770 \approx 0.245$$

这就是销售第二个工厂电灯泡的条件概率（假定电灯泡都符合标准）。从前面给出的式子里可以得出

$$P(\overline{B}) = 0.3 \qquad P_A(\overline{B}) \approx 0.245$$

（$\overline{B}$表示事件B的非）

例2 对某地区多年的观察表明，10万名年满十岁的儿童中，活到四十岁的平均82277人，活到七十岁的37977人。求年满四十岁的人能活到七十岁的概率。

因为年满四十岁的82277人中活到七十岁的平均有37977人，则年满四十岁的人活到七十岁的概率等于

$$37977/82277 \approx 0.46$$

如果用A和B表示下列事件：A——十岁儿童活到了七

㊀ 这用下列方法易于算出。每1000个电灯泡中平均有700个是第一个工厂生产的，而第一个工厂每100个电灯泡中平均有83个质量符合标准。因此第一个工厂700个电灯泡中，质量符合标准的平均有7×83＝581个。其余189个质量符合标准的电灯泡是第二个工厂生产的。

十岁，B ——四十岁人活到了七十岁，则显然有

$$P(A)=0.37977\approx0.38 \qquad P_B(A)\approx0.46$$

§8 概率的乘法定理

再来研究一下上一节的例1。在每一千个销售的电灯泡中，第二个工厂生产的平均为300个，而这300个中质量符合标准的平均有189个。由此得到电灯泡为第二个工厂生产（事件$\overline{B}$）的概率是

$$P(\overline{B})=300/1000=0.3$$

而在第二个工厂生产的这一条件下，质量符合标准的概率为

$$P_{\overline{B}}(A)=189/300=0.63$$

因为每一千个电灯泡中，第二个工厂生产的而且符合质量标准的有189个，则事件A与$\overline{B}$同时出现的概率等于

$$P(A\text{与}\overline{B})=\frac{189}{1000}=\frac{300}{1000}\times\frac{189}{300}=P(\overline{B})\cdot P_{\overline{B}}(A)$$

这个乘法定理可方便地推广到一般情况。假设在每批n次作业中，结果B平均出现m次，而在每批结果B出现的m次作业中，结果A出现l次，因此

$$P(B)=m/n \qquad P_B(A)=l/m$$

$$P(A\text{与}B)=\frac{l}{n}=\frac{m}{n}\cdot\frac{l}{m}=P(B)\cdot P_B(A) \qquad (1)$$

乘法定理　两个事件同时出现的概率等于第一个事件的概率与假定第一个事件业已存在的条件下所算出的第二个事件的条件概率相乘之积。

当然，可以把两个事件中的任一个称为第一个事件，因

此根据式（1），可以同样地写出：

$$P(A与B) = P(A)\cdot P_A(B) \qquad (1')$$

由此可得出重要的关系式：

$$P(A)\cdot P_A(B) = P(B)\cdot P_B(A) \qquad (2)$$

在前面所举的例子里，有

$$P(A与\overline{B}) = 189/1000 \qquad P(A) = 77/100$$

$$P_A(\overline{B}) = 189/770$$

由此证明式（1′）是正确的。

例 在某一工厂里有96%的产品合格（事件A）；每一百个合格产品中平均有75个是一等品（事件B）。求该工厂生产的一等产品的概率。

因为要使产品是第一流的，就须使产品是合格的（事件A）且又是一等品（事件B），所以求P（A与B）。

由问题的条件可知

$$P(A) = 0.96 \qquad P_A(B) = 0.75$$

因此根据公式（1′）

$$P(A与B) = 0.96\times 0.75 = 0.72$$

§9 独立事件

不同纺纱机生产的两卷纱的强度试验表明，某一长度的第一卷纱的试样经得住标准载荷的概率为0.84㊀，而第二卷纱——概率为0.78。求取自两个不同纱卷的两种纱线试样都处在经受住标准载荷状态的概率。

㊀ 这意味着从第一卷纱取出的100件试样中，平均有84件试样经得住这一载荷，而16件试样经不住并且断了。

用A表示从第一卷纱中取出的试样经受住标准载荷的事件，同样用B表示从第二卷纱中取出的试样经受住标准载荷的事件。因为是求P(A与B)，则采用乘法定理：

$$P(A与B) = P(A)P_A(B)$$

显然式中$P(A)=0.84$，但$P_A(B)$等于多少呢？根据条件概率的一般定义，$P_A(B)$是对于第一卷纱试样经得住的标准载荷 第二卷纱试样同样也能经得住的概率。不过，虽然我们可以同时进行这些试验，但是纱线试样取自各纺织机生产的完全不同的纱卷，因此事件B的概率不取决于事件A是否发生。这实际意味着，第二卷纱试样经得住标准载荷试验的百分比不取决于第一卷纱试样的结实程度，即

$$P_A(B) = P(B) = 0.78$$

由此得出

$$P(A与B) = P(A)\cdot P(B) = 0.84\times 0.78 = 0.6552$$

正如我们所见，这个例子与前面各个例子的不同之处在于：尽管给总条件附加了事件A存在这一条件，但B的概率并不随之改变。换句话说，条件概率$P_A(B)$等于无条件概率$P(B)$。在这种情况下认为**事件B与事件A无关**。

不难相信，如果B与A无关，那么A也与B无关。其实，如果$P_A(B)=P(B)$，则根据公式(2)，有$P_B(A)=P(A)$，而这就表示事件A与事件B无关。这样，两个事件的独立性是相互的。由此可知，对于互为独立的事件，乘法定理可变成特别简单的形式：

$$P(A与B) = P(A)P(B) \qquad (3)$$

就象在应用加法定理的所有情况下必须预先规定各个事件之间互不相容性一样，在应用公式(3)的所有情况下，

必须肯定事件A与B相互独立。忽视这一条就会酿成大错。如果事件A与B相互关联，则公式(3)不正确，在这种情况下应该用较为普遍的公式(1)或(1′)。

将公式(3)推广来求三个或多个(而非两个)相互独立事件出现的概率甚为容易。例如，有三个相互独立的事件A、B、C（这意味着，其中任一事件的概率不依赖于其他两个事件出现与否）。因为事件A、B、C相互独立，则根据公式(3)

$$P(A与B与C)=P(A与B)\cdot P(C)$$

将公式(3)代入，得到

$$P(A与B与C)=P(A)P(B)P(C) \qquad (4)$$

显然，只有当所研究的一组任意个事件相互独立(即其中每个事件的概率不依赖于其余事件出现与否）的情况下，上式才能成立。

任意个相互独立的事件同时出现的概率等于这些事件的概率的乘积。

例1 某工人看三台纺织机，在一小时内纺织机不要求工人维护的概率是，第一台为0.9，第二台为0.8,第三台为0.85。求一小时内一台纺织机也不需要工人维护的概率。

由于纺织机的运转相互独立，根据公式（4）所求的概率等于

$$0.9\times0.8\times0.85=0.612$$

例2 在例1的条件下，求在一小时内三台纺织机中至少有一台不要求工人维护的概率。

由于这里所求的概率是P（A，或B，或C）型的，因此首先想到加法定理，但是很快考虑到这个定理在这里不适

用，因为三个条件中的任意二个都是彼此相容的（即在同一时间里，任何一台机器不妨碍另两台机器顺利地工作），而且很快发现加法定理与此处讨论的问题无关，因为三个已知概率之和大大超过1，因此这不可能是任何一个的概率。

同时也注意到纺织机需要工人维护的概率：第一台为0.1，第二台为0.2，第三台为0.15。因为这三个事件相互独立，则按公式（2），三个事件都出现的概率为

$$0.1\times0.2\times0.15=0.003$$

但是，“三台纺织机都要求工人维护”的事件和“三台纺织机中至少有一台不要求工人维护”的事件显然是一对对立事件。因此，它们概率之和等于1。因此所求的概率为

$$1-0.003=0.997$$

当事件的概率如此接近于1时，这个事件实际上可认为是必然事件。这意味着，在一小时内三台纺织机中总是至少有一台在顺利地工作，而不要求工人维护。

例3 在一定的条件下，在试验台上试验250台仪器。在一小时内，其中某台仪器发生故障的概率等于0.004，而且这一概率对所有仪器都是一样的。求在一小时内至少有一台仪器出故障的概率。

一台仪器不出故障的概率应是

$$1-0.004=0.996$$

根据独立事件的乘法定理，250台仪器中一台都不出故障的概率等于250个因子（每个因子为0.996）的乘积，即等于$(0.996)^{250}$。于是仪器中至少有一台发生故障的概率等于

$$1-(0.996)^{250}$$

这里不作详细计算。计算结果表明，这个数近似等于5/8。

因此，在一小时内，每台仪器发生故障的概率尽管很小，但是，在大量仪器进行试验时，至少有一台仪器出现故障的概率却相当可观了。

由上述二例很容易得出一个重要的规则。即一些互为独立事件A_1，A_2、……、A_n中至少有一个事件出现的概率$P(A_1,或A_2,\cdots,或A_n)$。如果用$\overline{A}_K$表示A_K未发生的事件，则事件A_K与$\overline{A}_K$是相互对立的事件，因此

$$P(A_K)+P(\overline{A}_K)=1$$

另一方面，事件$\overline{A}_1$、$\overline{A}_2$、……、$\overline{A}_n$显然是相互对立的；结果

$$\begin{aligned}&P(\overline{A}_1,与\overline{A}_2,与\cdots,与\overline{A}_n)\\&=P(\overline{A}_1)P(\overline{A}_2)\cdots P(\overline{A}_n)\\&=[1-P(A_1)][1-P(A_2)]\cdots[1-P(A_n)]\end{aligned}$$

最后，事件$(A_1,或A_2,\cdots,或A_n)$与$(\overline{A}_1,与\overline{A}_2,\cdots,与\overline{A}_n)$又是互为独立的事件(或者至少有一个事件$A_K$出现，或者所有事件$\overline{A}_K$都出现，因此，

$$\begin{aligned}&P(A_1,或A_2,\cdots,或A_n)\\&=1-P(\overline{A}_1,与\overline{A}_2,\cdots,与\overline{A}_n)\\&=1-[1-P(A_1)][1-P(A_2)]\cdots[1-P(A_n)]\end{aligned}\qquad(5)$$

这个重要的公式可以计算出事件A_1、A_2……、A_n在其给定的概率之下，**至少有一个**事件出现的概率。但是，这个公式仅在这些事件是互为独立事件的情况下才能成立。在特殊

情况下，即当所有事件A_K具有同一概率值p时（如例3），则

$$P(A_1, 或A_2, \cdots, 或A_n) = 1-(1-p)^n \quad (6)$$

例4 加工一批矩形六面体仪器零件。如果每边边长偏离规定尺寸不大于0.01mm则认为零件合格。偏差超过0.01mm的概率是：

长 $p_1=0.8$

宽 $p_2=0.12$

高 $p_3=0.1$

求零件不合格的概率P。

若零件不合格，则它至少有一条边长与规定尺寸的偏差大于0.01mm。通常这三个事件可认为是相互独立的（因为这些偏差基本上是由各种原因引起的）。因此解此题可用公式（5）。于是

$$P=1-(1-p_1)(1-p_2)(1-p_3)\approx 0.27$$

由此可知，100个零件中平均有73个合格。

第四章　加法定理与乘法定理的推论

§10　若干不等式的导出

现重新回顾前一章关于电灯泡的例子。先引入以下表示事件的符号：

A——符合质量标准的电灯泡；

$\overline{A}$——不符合质量标准的电灯泡；

B——第一个工厂生产的电灯泡；

$\overline{B}$——第二个工厂生产的电灯泡。

显然，事件A与事件$\overline{A}$是一对对立事件；事件B与$\overline{B}$也是一对对立事件。如果电灯泡符合标准（事件A），则它或者是第一个工厂生产的（事件A与事件B），或者是第二个工厂生产的（事件A与事件$\overline{B}$）。因为这两个事件互不相容，根据加法定理：

$$P(A)=P(A与B)+P(A与\overline{B}) \quad (1)$$

同理可得

$$P(B)=P(A与B)+P(\overline{A}与B) \quad (2)$$

再来研究事件（A或B），会出现三种可能性：

1）　A与B；

2） A与$\overline{B}$；

3） $\overline{A}$与B。

三种可能性中任意二种是互不相容的，因此根据加法定理

$$P(A或B)=P(A与B)+P(A与\overline{B})+P(\overline{A}与B) \quad (3)$$

将（1）式与（2）式相加并考虑到（3）式，很容易得到：

$$P(A)+P(B)=P(A与B)+P(A或B)$$

于是

$$P(A或B)=P(A)+P(B)-P(A与B) \quad (4)$$

现在得到了一个很重要的结果。此处虽是讨论一个具体实例，但它具有普遍意义，其结论可适用于任意一对事件A和B。在这之前，概率P(A或B)表达式是在对事件A、B给予特定假设条件下得到的（开始，假设它们是不相容的，然后，假设它们是相互独立的）。现在得到了对于任意一对事件A和B，在无任何补充假设条件下都成立的公式（4）。

我们应当注意公式（4）与前面导出的一些公式之间的一个实质性区别。在以前的公式里，概率P(A或B)总是只用概率P(A)和P(B)表示，因此，知道事件A与事件B的概率就能够求出（A或B）的概率。在公式（4）里情况就不同了。为了计算P(A或B)的值，除了知道P(A)和P(B)外，还必须知道P(A与B)，即事件A和B同时出现的概率。而且在一般情况下，即在事件A与事件B之间为任意关系之下，求概率P(A与B)并不比求概率P(A或

B）容易。因此，人们很少采用公式（4）进行直接的实际计算，但是它却有很大的理论价值。首先我们深信，作为个别例子，从公式（4）很容易求出以前的各个公式。如果事件A与事件B彼此不相容，则事件（A与B）的概率不可能出现。即P(A与B）=0。公式（4）就变为下式：

$$P（A或B）=P（A）+P（B）$$

即变为加法定理表达式。如果事件A与事件B相互独立，则根据第23页公式（3）

$$P(A与B）=P(A）P(B)$$

由公式（4）得到

$$P(A或B）=P(A）+P(B）-P(A）P(B)$$
$$=1-[1-P(A）][1-P(B）]$$

这就是第26页的公式（5）（$n=2$的情况）。

此外，从公式（4）可得出一个重要的推论。因为在所有情况下，P(A与B）≥0，那么从公式（4）得出在所有情况下

$$P(A或B）\leqslant P(A）+P(B) \qquad (5)$$

该不等式很容易推论到任意数量的事件。例如，在三个事件的情况下，根据公式（5），

$$P(A，或B，或C）\leqslant P(A或B）+P(C)$$
$$\leqslant P(A）+P(B）+P(C)$$

显然用同样的方法可从三个事件引伸到四个事件。依此类推，得到了普遍性的结论：

在若干事件中至少出现一个事件的概率总是小于这些事件的概率之和。这里，只有当这些事件中任意两个事件都互不相容时，上式两边才相等。

§11 全概率公式

我们再来讨论第18页关于电灯泡的例子，并仍采用第28页引入的几种符号。正如前面所述，第二个工厂生产的符合质量标准的电灯泡的概率为

$$P_{\bar{B}}(A)=\frac{189}{300}=0.63$$

第一个工厂生产的符合质量标准的电灯泡的概率为

$$P_B(A)=\frac{581}{700}=0.83$$

假设我们不仅知道这两个数字，并知道电灯泡在第一个工厂生产的概率为

$$P(B)=0.7$$

在第二个工厂生产的概率为

$$P(\overline{B})=0.3$$

求无条件概率P(A)，即求对电灯泡生产厂家不作任何假定的条件下，每个电灯泡符合标准的概率。

为解此题，我们作如下研究。用E表示电灯泡是第一个工厂生产的，且其质量符合标准这样一种双重事件；同样用F表示电灯泡是第二个工厂生产的，且其质量符合标准的双重事件。因为所有标准电灯泡是由第一个工厂或第二个工厂生产的，那么，事件A等于事件“E或F”，又因为事件E和事件F互不相容，根据加法定理，有

$$P(A)=P(E)+P(F) \tag{6}$$

从另一方面，要使事件E成立，就应该：1）电灯泡由

第一个工厂生产（B）；2）它符合标准（A）。因此，事件E等于事件（B与A）。由此，根据乘法定理

$$P(E)=P(B)P_B(A)$$

用完全相同的方法可得到：

$$P(F)=P(\overline{B})P_{\overline{B}}(A)$$

将它们代入等式（6）：

$$P(A)=P(B)P_B(A)+P(\overline{B})P_{\overline{B}}(A)$$

这个公式解决了求解问题，代入已知数据，求出$P(A)=0.77$。

举例 为播种而购买了一些甲等小麦种子，其中混有少量乙、丙、丁等种子。从这些种子中任意取一粒，这个事件是：该粒种子是甲等的用A_1表示，是乙等的用A_2表示，是丙等的用A_3表示，最后，是丁等的用A_4表示。已知任意取的一粒种子是这等或那等的概率为

$$P(A_1)=0.96,\ P(A_2)=0.01,$$
$$P(A_3)=0.02,\ P(A_4)=0.01$$

（这四个数之和等于1，对全事件系统而言，这是理所当然的）。

麦种长成每穗至少有50粒籽的麦穗的概率是：

1）甲等种子的为0.50；

2）乙等种子的为0.15；

3）丙等种子的为0.20；

4）丁等种子的为0.05。

求麦穗至少有50粒籽的无条件概率。

令K表示麦穗至少包含50粒籽的事件，则根据问题的条

$P_{A_1}(K)=0.50$　　$P_{A_2}(K)=0.15$

$P_{A_3}(K)=0.20$　　$P_{A_4}(K)=0.05$

我们的任务是求P(K)。用E_1表示甲等麦种，并且其生长麦穗至少包含50粒籽的事件，因而E_1与事件(A₁与K相等；用同样方式可以表示出：

E_2——事件（A_2与K）；

E_3——事件（A_3与K）；

E_4——事件（A_4与K）。

要使事件K出现，显然必须使事件E_1，E_2，E_3，E_4中的一个出现，由于这些事件中任意两个互不相容，则根据加法定理，我们得到

$$P(K)=P(E_1)+P(E_2)+P(E_3)+P(E_4) \tag{7}$$

从另一方面，根据乘法定理

$$P(E_1)=P(A_1与K)=P(A_1)\,P_{A_1}(K)$$

$$P(E_2)=P(A_2与K)=P(A_2)\,P_{A_2}(K)$$

$$P(E_3)=P(A_3与K)=P(A_3)\,P_{A_3}(K)$$

$$P(E_4)=P(A_4与K)=P(A_4)\,P_{A_4}(K)$$

将这些式子代入式（7），得到：

$$P(K)=P(A_1)\,P_{A_1}(K)+P(A_2)\,P_{A_2}(K)+P(A_3)\,P_{A_3}(K)+P(A_4)\,P_{A_4}(K)$$

显然，问题由此可解。

把已知数代入上式，求出

$$P(K)=0.486$$

根据上面两个类似的例子，得出一个重要的通用定理现在，可以毫不费力地用公式表示并给以证明。设某个作业有A_1、A_2、……A_n个结果，它们组成一个全事件系统（这意味着，这些事件中的任何两个事件互不相容，并且，其中必有某个事件要出现）。因此，对于该作业的任意一种可能的结果K，有下式成立：

$$P(K)=P(A_1)P_{A_1}(K)+P(A_2)P_{A_2}(K)+\cdots+P(A_n)P_{A_n}(K) \quad (8)$$

公式（8）通常称为**全概率公式**。采用与上两例中一模一样的方法就可证明该公式。首先，事件K的出现要求事件（Ai与K）中的某一个事件出现，因此根据加法定理

$$P(K)=\sum_{i=1}^{n}P(A_i与K) \quad (9)$$

其次，根据乘法定理

$$P(A_i与K)=P(A_i)P_{A_i}(K)$$

将此式代入公式（9），就得到公式（8）。

§12 巴叶斯公式

上一节的一些公式加上其他附加条件，可推导出一个重要结论。现从这个结论入手，在研究实例的基础上再阐述公式的现实意义。

再一次设事件A_1、A_2……、A_n是某个作业的全部结果。如果K表示该作业的任一结果，根据乘法定理

$$P(A_i与K)=P(A_i)P_{A_i}(K)=P(K)P_K(A_i) \quad (1\leqslant i\leqslant n)$$

由此

$$P_K(A_i)=\frac{P(A_i)\ P_{A_i}(k)}{P\ (K)} \qquad (1\leqslant i\leqslant n)$$

或者用全概率公式（8）表示上面分数的分母，得

$$P_K\ (A_i)\ =\frac{P(A_i)\ P_{A_i}\ (K)}{\sum_{r=1}^{n}P(A_r)\ \ P_{A_r}(K)} \quad (1\leqslant i\leqslant n)$$

(10)

这就是**巴叶斯**公式，它在实际概率计算中有许多附加条件。在下图所示的各种情况下多半采用此公式。

假设对在直线MN上的目标进行射击（图3）。设想把该直线分为五段：a、b′、b″、c′、c″。并且不知道目标的准确位置，只知道目标在五段中这一段或那一段的概率，并假设这些概率为

$P(a)\ =0.48 \qquad P(b')\ =P(b'')\ =0.21$

$P(c')\ =P(c'')\ =0.05$

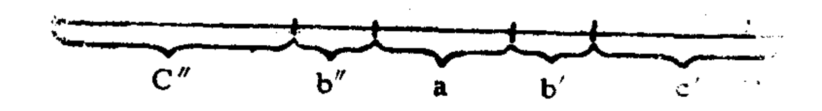


图 3

这些概率的和等于1。

用a、b′、b″、c′、c″分别表示目标位于各线段上的事件。因为知道目标在线段a上的概率最大，所以，射手自然朝它射击的次数多。由于不可避免的射击误差，也可能偏离

a段，而在其它段上命中目标。设命中目标（事件K）的概率是：

如果目标位于线段a上，则$P_a(K)=0.56$

如果目标位于线段b′上，则$P_{b'}(K)=0.18$

如果目标位于线段b″上，则$P_{b''}(K)=0.16$

如果目标位于线段c′上，则$P_{c'}(K)=0.06$

如果目标位于线段c″上，则$P_{c''}(K)=0.02$

假定射击后命中了目标，则构成事件K。目标在各线段上的概率——前面已有的（即P(a)，P(b′)……）数值——要重新估计。这种重新估计由定性判断就一目了然，无需计算，当朝a段射击并命中目标时，概率P(a)显然增大。但是，我们想准确地、定量地求出重新估计值，即想求目标在各种可能线段上被命中的概率$P_K(a)$，$P_K(b')$……的准确表达式。巴叶斯公式可以立即求解这一问题：

$$P_K(a)=P(a)P_a(K)[P(a)P_a(K)+P(b')P_{b'}(K)+P(b'')P_{b''}(K)+P(c')P_{c'}(K)+P(c'')P_{c''}(K)]^{-1}\approx0.8$$

由此看到，$P_K(a)$的确比P(a)大。

同理，可以容易地求出目标在其他线段上被命中的概率$P_K(b')$……。

由巴叶斯公式给出的上述概率表达式彼此之间的区别只是分子不同，而它们的分母相同，即P(K)≈0.34。指出这一点对计算概率是有益的。

解同类问题的一般思路可作如下描述。设某种作业的条件包含某个未知项，关于它可以有n个不同的假设事件，即A_1、A_2、…A_n，它们组成事件的全系统。由于这种或那种原

因，在试验之前已经知道了这些假设事件的概率$P(A_i)$，也知道假设的A_i"联系"某一事件K（例如，命中目标）的概率$P_{A_i}(K)$ $(1\leqslant i\leqslant n)$ （$P_{A_i}(K)$是在假设的A_i为真的条件下事件K的概率）。如果试验结果事件K出现。则势必要重新估计假设事件A_i的概率，并求这些假设事件的新概率$P_K(A_i)$。用巴叶斯公式即可解这个问题。

在炮兵射击演习中有一种所谓"试射"射击，其目的是搞准射击条件。这时，需搞清楚的未知项不仅有目标的位置，而且有影响射击效率的其他任一项射击条件（其中包括所用武器的这种或那种特性）。

这种射击往往不止一次而是多次，并且根据已得到的射击结果提出计算新的假设概率的问题。这种情况屡见不鲜。在所有这些情况下，提出的问题可用巴叶斯公式迎刃而解。

为简化起见，在上述一般公式里，令

$$P(A_i)=P_i,\quad P_{A_i}(K)=p_i \qquad (1\leqslant i\leqslant n)$$

于是巴叶斯公式变成简单形式

$$P_K(A_i)=\frac{P_i p_i}{\sum_{r=1}^{n} P_r p_r}$$

假定进行了s次试射，事件K出现m次，K未出现次数为s—m。用K^*表示s次射击中得到的结果。又假定，每次射击结果是相互独立的。如果A_i这一假设为真，结果K的概率等于p_i，那么，就意味着对立事件（即K未出现的事件）的概率等于$1-p_i$。

根据独立事件的乘法定理，在射击中事件K出现确定的

m次的概率等于$p_i^m(1-p_i)^{s-m}$。由于事件K出现m次的射击可能是s次射击中的任意m次，则事件K^*可能为C_s^m个不相容事件。这样一来，根据概率的加法定理

$$P_{A_i}(K^*)=C_s^m p_i^m(1-p_i)^{s-m}\quad(1\leqslant i\leqslant n)$$

于是，巴叶斯公式可为

$$P_{K^*}(A_i)=\frac{P_i p_i^m(1-p_i)^{s-m}}{\sum_{r=1}^{n}P_r p_r^m(1-p_r)^{s-m}}\quad(1\leqslant i\leqslant n)\qquad(11)$$

这就解决了前面提出的问题。显然，类似的问题不仅发生在炮兵演习之中，而且存在于人类活动的其他一些领域。

例1 在本节开头所研究的问题里，如果对a段的两次连续射击都命中了目标，求目标在该段的概率。

用K^*表示两次命中目标，根据公式（11）有

$$P_{K^*}(a)=\frac{P(a)[P_a(K)]^2}{P(a)[P_a(K)]^2+P(b')[P_{b'}(K)]^2+\cdots}$$

经过不甚复杂的计算后表明，由于两次命中的结果，目标位于线段a的概率还要增大。

例2 在某项生产中，产品符合标准的概率等于0.96。设有一简化的试验系统㊀，对于符合标准的产品，给出正结果的概率为0.98．对于不符合标准的产品，给出负结果的概率

㊀ 简化检验实际上非常必要。例如，电灯泡投放市场时，如果对其使用寿命都要进行点燃检验，而且点燃时间不少于1200小时，那么用户就会买到一个报废或即将报废的电灯泡。因此人们采用别的办法，比如用检查电灯泡的耐燃性试验来代替电灯泡的点燃时间试验。

只有0.05。求某种产品两次经受简化试验，符合标准的概率？

这里，全部假定的事件系统由两个对立事件构成：1）产品符合标准；2)产品不符合标准。在试验之前，这些假定事件的概率分别等于$P_1=0.96$和$P_2=0.04$。在第一个假定事件里，产品经受住试验的概率$p_1=0.98$，而在第二个假定事件里，$p_2=0.05$。在两次实验之后，根据公式（11），第一个假定事件的概率等于

$$\frac{P_1p_1^2}{P_1p_1^2+P_2p_2^2}=\frac{0.96\times(0.98)^2}{0.96\times(0.98)^2+0.04\times(0.05)^2}$$

$$\approx 0.9999$$

我们看到，如果产品经受了本题所说的试验，则在一万件产品中只有一件产品可能不符合标准。当然，这完全满足实际要求。

例3 在研究某一病人症状时，怀疑他患有三种疾病中的一种：A_1、A_2、A_3。在给定条件下它们的概率是：

$$P_1=\frac{1}{2},\qquad P_2=\frac{1}{6},\qquad P_3=\frac{1}{3}$$

为搞清病情，决定作几次化验，化验后给出肯定结果的概率如下：疾病A_1为0.1，疾病A_2为0.2，疾病A_3为0.9，曾经作了五次化验，其中四次为肯定结果，一次为否定结果。现求化验后每种疾病的概率。

对于疾病A_1，按乘法定理，被指出的化验结果的概率为$p_1=C_5^4\times(0.1)^4\times0.9$。对第二种疾病而言，这种概率为$p_2=C_5^4\times(0.2)^4\times0.8$。而对于第三种疾病而言，

$p_3 = C_5^4 \times (0.9)^4 \times 0.1$。

根据巴叶斯公式，求出化验后疾病A_1的概率为

$$\frac{P_1p_1}{P_1p_1+P_2p_2+P_3p_3}$$

$$=\frac{\frac{1}{2}\times(0.1)^4\times0.9}{\frac{1}{2}\times(0.1)^4\times0.9+\frac{1}{6}\times(0.2)^4\times0.8+\frac{1}{3}\times(0.9)^4\times0.1}$$

≈ 0.002

疾病A_2的概率为

$$\frac{P_2p_2}{P_1p_1+P_2p_2+P_3p_3}$$

$$=\frac{\frac{1}{6}\times(0.2)^4\times0.8}{\frac{1}{2}\times(0.1)^4\times0.9+\frac{1}{6}\times(0.2)^4\times0.8+\frac{1}{3}\times(0.9)^4\times0.1}$$

≈ 0.01

疾病A_3的概率为

$$\frac{P_3p_3}{P_1p_1+P_2p_2+P_3p_3}$$

$$=\frac{\frac{1}{3}\times(0.9)^4\times0.1}{\frac{1}{2}\times(0.1)^4\times0.9+\frac{1}{6}\times(0.2)^4\times0.8+\frac{1}{4}\times(0.9)^4\times0.1}$$

≈ 0.988

因为这三个事件A_1、A_2、A_3在化验之后组成了全事件系统，那么若要检验计算结果，可把三个得数相加。不言而喻，其和必然等于1。

第五章 伯努利试验

§13 实　　例

例1 某一品种棉花纤维，其长度小于45mm的占75%，长度大于（或等于）45mm的占25%。求任取三条纤维，其中二条短于45mm，一条长于45mm的概率。

用A表示选出纤维长度小于45mm的事件，用B表示选出的纤维长度大于45mm的事件，则显然

$$P(A)=\frac{3}{4} \qquad P(B)=\frac{1}{4}$$

进一步约定用AAB方式表示复杂事件：选出的头两根纤维短于45mm，而第三根长于45mm。显然，也有BBA、ABA等方式，这里需要计算的是事件C的概率，即三条纤维中，二条短于45mm，一条长于45mm的概率。为此就应该出现下列方式中的一种，即

$$AAB,\ ABA,\ BAA \qquad (1)$$

因为这三种结果中的任意两种彼此是不相容的，则按加法定理

$$P(C)=P(AAB)+P(ABA)+P(BAA)$$

因为我们可以把选择纤维的结果看作相互独立的事件，所以上式右边三个被加数彼此相等。根据独立事件各概率相乘的乘法定理，（1）式每个方式的概率是三个乘数的乘积，其中二个是P(A)＝3/4，而另一个是P(B)＝1/4，这样，式（1）三个方式中每一个的概率为

$$\left(\frac{3}{4}\right)^2 \times\frac{1}{4}=\frac{9}{64}$$

于是，

$$P\ (C)\ =3\times\frac{9}{64}=\frac{27}{64}$$

例2 经过数十年连续观察，人们发现，每一千名新生儿中平均有515个男孩和485个女孩。在某一有六个孩子的家庭里，求女孩不多于两个的概率。

要使我们所求的事件出现，在这个家庭里女孩子数就应该是0或1，或2。我们用P_0、P_1、P_2相应地表达这些局部事件的概率。显然，根据概率的加法定理，所求概率为

$$P=P_0+P_1+P_2 \qquad (2)$$

对每个孩子而言，男孩的概率等于0.515，女孩的概率等于0.485。

首先求P_0，它是该家庭中所有孩子都是男孩的概率。因为可以把生这种或那种性别孩子的事件看成是同生其他几个孩子无关的独立事件。则根据乘法定理，所有六个孩子都是男孩的概率等于六个乘数（等于0.515）的乘积，即

$$P_0=\ (0.515)^6\approx0.018$$

再计算概率P_1，即家里六个孩子中，一个是女孩，其余五个是男孩的概率。由于女孩可能是出生顺序中的某个孩子（第一个，第二个等等），因此该事件可能以六个不同的方式发生。我们来研究以其中任一种方式出现的事件，例如，第四个孩子是女孩。这种方式的概率按乘法定理等于六个乘数之积，它们之中五个乘数是0.515，第六个乘数（占据第四个位置的）是0.485，即该概率为$(0.515)^5\times0.485$。

这也是五种可能方式中任一方式的概率。因此，根据加法定理该事件的概率P_1等于六个$(0.515)^5 \times 0.485$数之和，即

$$P_1 = 6 \times (0.515)^5 \times 0.485 \approx 0.105$$

现在来求P_2——即两个孩子是女孩,四个孩子是男孩的概率。和前面一样,我们立即发现,该事件能以一系列方式发生（例如，其中一种方式是：按出生顺序，第二个、第五个孩子是女孩，而其余的是男孩）。根据乘法定理，每种方式的概率等于$(0.515)^4 \times (0.485)^2$。因此根据加法定理，$P_2$等于乘数$(0.515)^4 \times (0.485)^2$乘以被研究事件的所有方式数；于是，全部问题就归结为求这最后一个方式数了。

每种方式都表示六个孩子中，两个是女孩，其余是男孩的情况。因此，各种各样方式的数目等于从六个孩子中选择两个孩子的不同选择法的数目，这种选择方法数等于从六个中选取两个的组合数，即

$$C_6^2 = \frac{6 \times 5}{2 \times 1} = 15$$

这样一来，

$$P_2 = C_6^2 \times (0.515)^4 \times (0.485)^2$$
$$= 15 \times (0.515)^4 \times (0.485)^2 \approx 0.247$$

将所得结果相加，得到

$$P = P_0 + P_1 + P_2 \approx 0.018 + 0.105 + 0.247 = 0.370$$

因而，在这样多子女的家庭中，女孩数小于1/3（六个中不多于两个），而男孩大于2/3的概率接近十分之四（P=0.37）。

§一4 伯努利公式

在上一节通过一系列例子了解到，每次都可能出现某一事件A的重复试验。"试验"一词包括非常丰富的、形形色色的内容。例如，如果对某一目标进行射击，那么，每次单独射击就意味着一次试验。如果检测电灯泡的燃烧寿命。则把每个灯泡的检测称为一次试验。如果按性别、体重或身高研究新生儿的组成，则调查每个婴儿就是一次试验。一般地说，进而可以把某些条件（在这些条件下出现我们关心的某个事件）的实现也理解为一次试验。

现在，开始研究概率论中的一个主要公式，该公式除了在各个知识领域具有实用价值外，对数学科学分支——概率论本身也具有很大意义。这个公式就是研究各个相互独立试验的顺序性。所谓独立试验是指每个试验的这种或那种结果的概率与某些结果出现的或者在别的试验中出现的情况无关。在这些试验的每个试验之中，某一事件A可能出现（或不出现）的概率p与试验的号数无关。所描述的公式称为伯努利公式：该项研究的奠基人是17世纪瑞士著名学者雅可夫·伯努利。

我们已经在一些例子里碰到过伯努利公式，关于这一点只要回顾上一节的例子即可信服。现在来求解下面的一般性问题。本章已研究过的例子都是它的局部情况。

问题 在一定条件下，在每个试验中，事件A出现的概率等于p。求一组n个独立试验中事件A出现k次，未出现=k次的概率。

把需要求概率的事件分成一系列方式，为了得到一种确定的方式，我们当以任意形式从一组试验中选取k个试验，并假定，在这k个试验中事件A出现，而在其余n－k个试验中事件A未出现。这样一来，每一方式要求出现n个确定的结果，其中事件A出现k次，未出现n－k次，根据乘法定理得到，每一确定方式的概率等于

$$p^k(1-p)^{n-k}$$

各种可能的方式数等于由n个试验组成的每一组中取k个试验的不同组数，即等于C_n^k。利用加法定理及已知的组合数学公式得

$$C_n^k=\frac{n(n-1)\cdots[n-(k-1)]}{k(k-1)\cdots 2\cdot 1}$$

在n个独立试验中，事件A出现k次的概率等于

$$P_n(k)=\frac{n(n-1)\cdots[n-(k-1)]}{k(k-1)\cdots 2\cdot 1}p^k(1-p)^{n-k} \qquad (3)$$

这就是所求之解。通常把C_n^k表达式变为另一种形式使用起来较为方便。将分子、分母都乘上$(n-k)[n-(k+1)]\cdots 2\cdot 1$这个因子，得

$$C_n^k=\frac{n(n-1)\cdots 2\cdot 1}{k(k-1)\cdots 2\cdot 1\cdot(n-k)[n-(k+1)]\cdots 2\cdot 1}$$

或者为简便起见，用阶乘n！表示从1到n的所有整数的乘积

$$C_n^k=\frac{n!}{k!\,(n-k)!}$$

于是$P_n(k)$可表示为

$$P_n(k)=\frac{n!}{k!\,(n-k)!}p^k(1-p)^{n-k} \qquad (4)$$

公式（3）和（4）通常称为**伯努利公式。**

当n和k为很大的数时，利用上式计算$P_n(k)$就相当困难，这是因为n!，k!，（n=k)!是很大的冗长数字。因此计算这类问题时，普遍采用专门编制的阶乘计算表，或采用某些近似公式。

实例 某企业用水量为正常值（每昼夜不大于规定的公升数）的概率等于3/4。求在1、2、3、4、5、6日连续六天内用水量为正常值的概率？

用$P_6(k)$表示六天中有k天用水量为正常值的概率，按公式（3）′求解（式中p＝3/4）

$$P_6(6)=\left(\frac{3}{4}\right)^6=\frac{3^6}{4^6}$$

$$P_6(5)=6\times(\frac{3}{4})^5\times\frac{1}{4}=\frac{6\times3^5}{4^6}$$

$$P_6(4)=C_6^4\times(\frac{3}{4})^4\times(\frac{1}{4})^2=C_6^2\ \frac{3^4}{4^6}=\frac{6\times5}{2\times1}\times\frac{3^4}{4^6}$$

$$=\frac{15\times3^4}{4^6}$$

$$P_6(3)=C_6^3\times(\frac{3}{4})^3\times(\frac{1}{4})^3=\frac{6\times5\times4}{3\times2\times1}\times\frac{3^3}{4^6}$$

$$=\frac{20\times3^3}{4^6}$$

$$P_6(2)=\frac{6\times5}{2\times1}\times(\frac{3}{4})^2\times(\frac{1}{4})^4=\frac{15\times3^2}{4^6}$$

$$P_6(1)=6\times\frac{3}{4}\times(\frac{1}{4})^5=\frac{6\times3}{4^6}$$

后，显然$P_6(0)$（六天中每天用水超量）为$1/4^6$。所有六个概率算式中，分母都相同，为$4^6=4096$。以上各式经约简计算得出

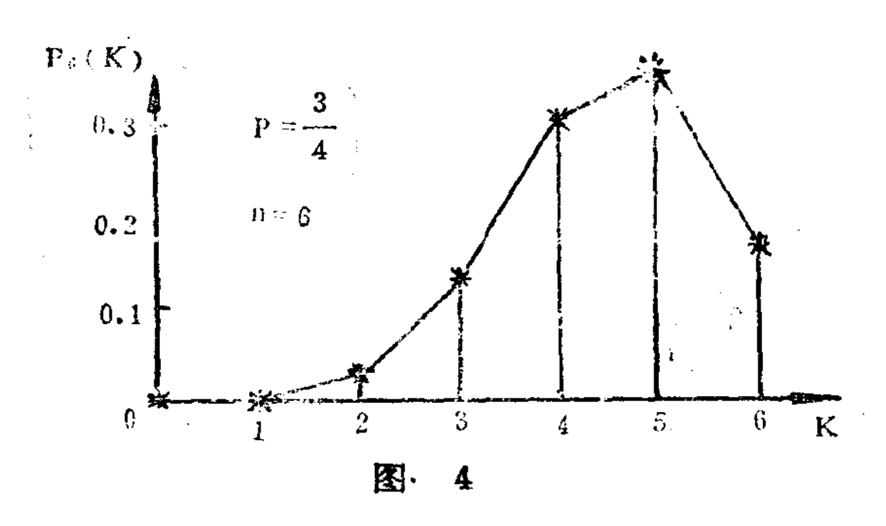
$P_6(6)\approx0.18$　　$P_6(5)\approx0.36$　　$P_6(4)\approx0.30$

$P_6(3)\approx0.13$　　$P_6(2)\approx0.03$　　$P_6(1)\approx P_6(0)\approx0$

由此可见，在六天中有一天或两天用水过量的概率最大，而在六天中有五天或六天用水过量的概率〔$P_6(1)+P_6(0)$〕实际上等于零。

§15　事件的最大概率对应値

上面研究的例子表明，在k日之内符合正常用水量的概率，起初随着k数的增加而增加，达到最大值后，开始减少；如果概率$P_6(k)$随数值k增加而变化的情况用图4的几何图形描述出来，则这种变化情况就一目了然了。



图·　4

当n的数目较大时，概率$P_n(k)$的数值随k的增加而变化的规律就会更加清楚。例如，当$n=15$，$p=1/2$时，变化

图形为图 5 所示。

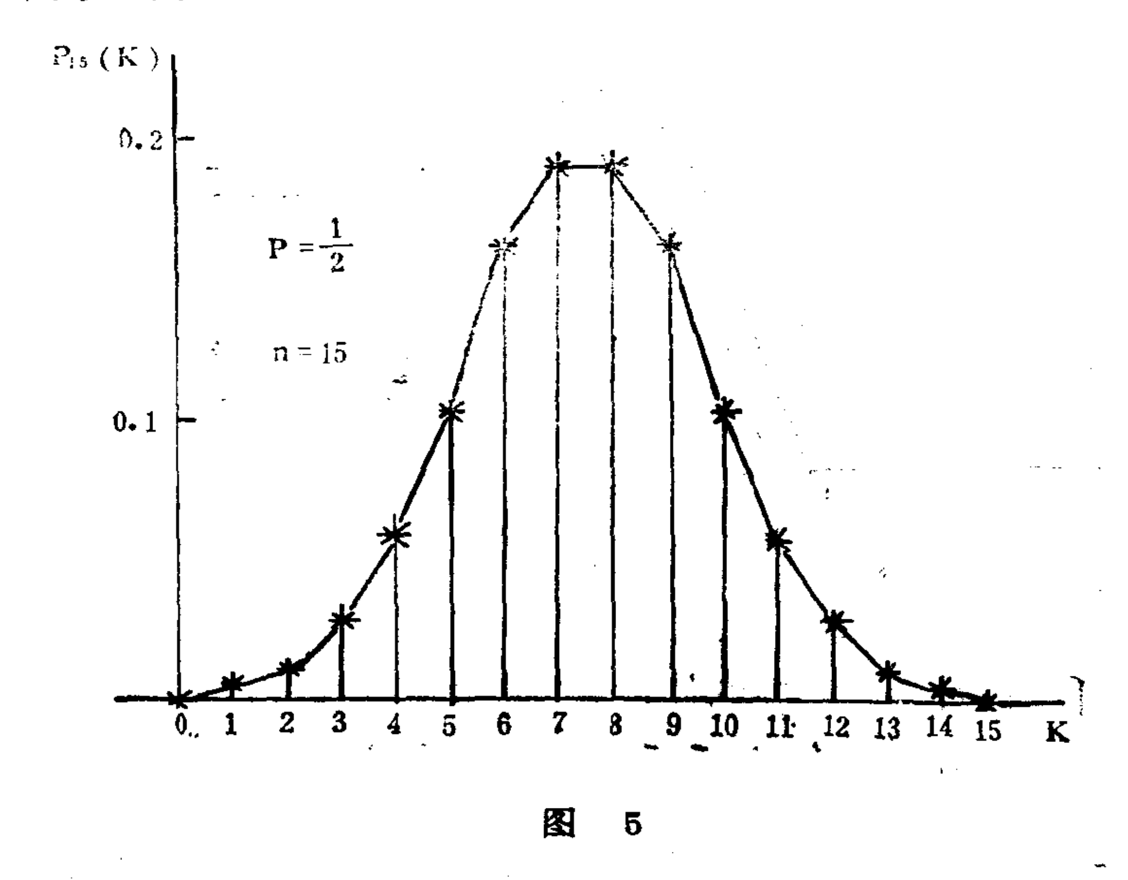


图 5

实际上有时要知道，所出现事件怎样的一个数值是对应于最大概率的，即在怎样一个k值时，概率$P_n(k)$为最大（当然在这种情况下，假定p与n是已知的）。

伯努利公式可以在一切情况下，简单地求解上述问题。我们现在就来研究这个问题。

首先计算 $\frac{P_n(k+1)}{P_n(k)}$ 的大小。根据公式（4）

$$P_n(k+1)=\frac{n!}{(k+1)!\ (n-k-1)}p^{k+1}(1-p)^{n-k-1} \tag{5}$$

并从公式（3）和公式（5）得到

$$\frac{P_n(k+1)}{P_n(k)}=\frac{n!\ k!\ (n-k)!\ P^{k+1}(1-p)^{n-k-1}}{(k+1)!\ (n-k-1)!\ n!\ p^k(1-p)^{n-k}}$$

$$=\frac{n-k}{k+1}\cdot\frac{p}{1-p}$$

概率$P_n(k+1)$是否大于、等于或小于概率$P_n(k)$，视$\frac{P_n(k+1)}{P_n(k)}$式是否大于、等于或小于1而定。这就归结为求以下三个关系式中哪一个能成立的问题。

$$\frac{n-k}{k+1}\cdot\frac{p}{1-p}>1,\quad \frac{n-k}{k+1}\cdot\frac{p}{1-p}=1,$$

$$\frac{n-k}{k+1}\cdot\frac{p}{1-p}<1 \qquad (6)$$

例如，如果想知道k为何值时能有$P_n(k+1)>P_n(k)$，则必须知道k为何值时有下面不等式成立

$$\frac{n-k}{k+1}\cdot\frac{p}{1-p}>1,\ 或\ (n-k)p>(k+1)(1-P)$$

由此得到： $np-(1-p)>k$

于是，当k值增长未达到$np-(1-p)$值之前，恒有$P_n(k+1)>P_n(k)$成立，即随着k值的增长，概率$P_n(k)$总是增加。例如，在图5所对应的问题里，$p=1/2$，$np-(1-p)=7$，这就是说，只要$k<7$，即对于从0到6所有数字，不等式$P_n(k+1)>P_n(k)$成立，图5也证实了这一点。

采用完全相同的方法，从(6)式中的其他两式求出

当$k=np-(1-p)$时，$P_n(k+1)=P_n(k)$

当$k>np-(1-p)$时，$P_n(k+1)<P_n(k)$

这样一来，随着k值的增加，一旦超过$np-(1-p)$数限，概率$p_n(k)$就开始减少，并减少到$P_n(n)$。

这个结论首先表明，上述各例所研究的$P_n(k)$值随k值增加而变化的情况（开始增加，然后减少）是一种普遍规律，它在所有情况下都成立。尤其是，这个结论可使我们立即解决曾经提过的问题——求数值k的最大概率对应值。用k_0表示最大概率对应值。那么

$$P_n(k_0+1)\leqslant P_n(k_0)$$

由此，根据前面所述，有

$$k_0\geqslant np-(1-p)$$

从另一方面，要使

$$P_n(k_0-1)\leqslant P_n(k_0)$$

根据前面公式，则应有

$$k_0-1\leqslant np-(1-p)$$

或

$$k_0\leqslant np-(1-p)+1=np+p$$

于是，k的最大概率对应值k_0应该满足两个不等式，即

$$np-(1-p)\leqslant k_0\leqslant np+p \qquad (7)$$

k_0处在从$np-(1-p)$到$np+p$的这一区间，简单的计算表明，该区间大小为1。因此，如果该区间某一端的数值，比如说$np-(1-p)$不是整数，则在区间内必将有一个且只有一个整数，即k_0可明确地算出。我们应该把下面情况视为正常的：由于$p<1$，因此，只在特殊情况下$np-(1-p)$的数值才是整数。在这种特殊情况下，不等式(7)确定了k_0的两个值：$np-(1-p)$和$np+p$。它们彼此差值为1。这两个值就是最大概率的对应值。而两个概率彼此相等，并超过k

对应的其他所有概率。这种特殊的情况确实存在。例如，在图5所描绘的实例中，$n=15$，$p=1/2$，即意味着 $np-(1-p)=7$，$np+p=8$；数字7和8是事件出现最大概率所对应的k的数值。其概率彼此相等，每一个近似等于0.196（所有情况如图所示）。

例1 对某个地区多年观察的结果表明，在7月1日下雨的概率等于4/17。求最近50年内7月1日是下雨天的最大概率对应值。这里$n=50$，$p=4/17$，

$$np-(1-p)=50\times\frac{4}{17}-\frac{13}{17}=11$$

这个整数表明，我们碰到了特殊情况；数11和12是两个相等的最大概率对应的下雨年数。

例2 在一次物理实验中，对一种粒子进行观察。在一定条件下，一定时间间隔内平均出现60个粒子，每个粒子运动速度大于v_0的概率是0.7。在另一些条件下，在同一时间间隔内，平均只出现50个粒子，而每个粒子运动速度超过v_0的概率等于0.8。在哪种试验条件下，运动速度超过v_0的粒子的最大概率对应数要大？

对于第一个试验条件，$n=60$，$p=0.7$，$np-(1-p)=41.7$，$k_0=42$。

对于第二个试验条件，$n=50$，$p=0.8$，$np-(1-p)=39.8$，$k_0=40$。

我们看到，在第一种实验条件下，"快速"粒子最大概率对应数较大。

实际上，往往碰到数值n非常大的情况（大批量射击、大批量生产等等）。在这种情况下，只要概率p不是非常

小，乘积np将是一个很大的数。在 $np-(1-p)$ 与 $np+p$ 两个表达式之间存在着事件发生的最大概率对应数，因为上两式的第二部分（即p和1－p）小于1，则上两式的数值和它们之间的事件发生的最大概率对应数值接近于np。例如若在15 s内连通电话用户的概率等于0.74，则在进入电话站的每一千个呼叫中，可以把1000×0.74作为在15 s内接通用户的最大概率对应数。

对这个结论可以赋予更加准确的公式。如果k_0表示n次实验中，事件出现的最大概率对应数，则k_0/n是在同样实验条件下，事件出现的最大概率对应数的"比率"。从不等式（7）得到

$$p-\frac{1-p}{n}\leqslant\frac{k_0}{n}\leqslant p+\frac{p}{n}$$

设想在单个试验里事件出现的概率p不变的情况下，使试验次数n越来越大（当然，在这种情况下，出现最大概率的对应数k_0也将增大）。那么，上面不等式里，左右两边的分数$(1-p)/n$和p/n将变得越来越小。这意味着，当n为大数时，这不等式左右两边两个分数可以忽略不计，也即意味着k_0/n的比值等于p。于是，**当试验次数为一大数时，事件发生的最大概率对应值的实际比率等于单个试验下事件发生的概率。**

比如说，如果在某些测量中，单个测量误差在α与β之间的概率等于0.84。那么，在大量测量时，可预期误差情况在α与β之间的最大概率也约为84%。当然，这并不是说，得到84%的这类误差的概率数值很大，相反，当测量数为大数时，这种"最大概率"本身将很小（由图5可见，最大概

率为0.196，这里只有15次试验，当试验数目很大时，最大概率还要小得多）。这个概率称为最大概率只是比较而言。有84%的测量，其误差在α与β之间的概率比有83%或86%的测量，其误差在α与β之间的概率要大。

从另一方面来看更容易明白，在一长串测量中，某一量值的这种或那种个别误差数的概率并不令人感兴趣。例如，进行200次测量，那么，计算200次中有137次以给定的精度测量的概率，这未必是必要的。因为实际上，这个数可以是137，136或138，甚至可能是140。相反，比方说，在200次测量中，误差在给定范围之内的测量次数大于100的概率，或者100至125之间的概率，或者它小于50的概率，计算以上概率问题无疑是很令人感兴趣的。如何计算这类概率呢？比如要计算测量数在100和120（含120）之间概率。确切地说，求下列不等式的概率：

$$100<k\leqslant 120$$

其中，k—— 命中数。要使这个不等式成立，k应等于101，102，……120等20个数中之，根据加法定理，该概率为

$$P(100<k\leqslant 120)=P_{200}(101)+P_{200}(102)+\cdots+P_{200}(120)$$

要直接计算上式的和，首先根据公式(3)计算20个类型为$P_n(k)$的各个部分的概率。当这些数都是一些大数字时，这种计算是相当困难的。因此，上式的和数实际上永远不能直接计算出来。解此问题有一些简便的近似公式和图表。得出这些公式和图表是采用这里将不涉及的复杂的数学分析方法。但是，通过对求$P(100<k\leqslant 120)$类型的概率的简单讨论，可以得到在许多情况下求所提问题详解的方法。关于这个问题

将在下一章里讲述。

第六章　伯努利定理

§16　伯努利定理概述

我们再详细地研究一下图5。图中，事件出现不同次数k的概率$P_{15}(k)$用垂直线段表示。根据加法定理，任意k值区间的概率等于该区间所有垂直线段长度之和。显而易见，同样长度的不同区间的概率之和是完全不同的。例如，区间$2\leqslant k<5$和$7\leqslant k<10$，其长度相同，其中每个区间的概率用三个垂直线段之和来表示。

由图可知，第二区间对应的概率之和比第一区间的要大得多。现已知道，所有n值的概率$p_n(k)$的分布图基本上呈图5所示的形式，即$p_n(k)$的数值开始随着k的增加而增加，在达到最大值后，随着k值的增加而减少。因此，对于长度相同的两段k值，靠近最大概率对应值k_0的区间将总是具有较大的概率数值。而包含k_0的区间具有的概率值比同长度的其他区间都大。

看来，在这方面需要多说几句。在n次试验中，事件出现次数k的所有可能数为$n+1$（$0\leqslant k\leqslant n$）。

取包含k_0的一段，该区间只占数值n可能数值的很小一部分，例如百分之一。这时，如果试验次数n很大，那么，这一区间具有最大概率数值，而其他k值所具有的概率值非常之小。这样一来，虽然所取的区间比k值小得多（区间长度只占全部k值长度的百分之一），但是，分布在该区间的垂

直线段长度之和比其他垂直线段长度之和要大得多。原因是：图5上中央区间垂直线段的长度比边缘部分垂直线段的长度长好多倍。当n数较大时，$P_n(k)$ 的分布图形近似地如图6所示。

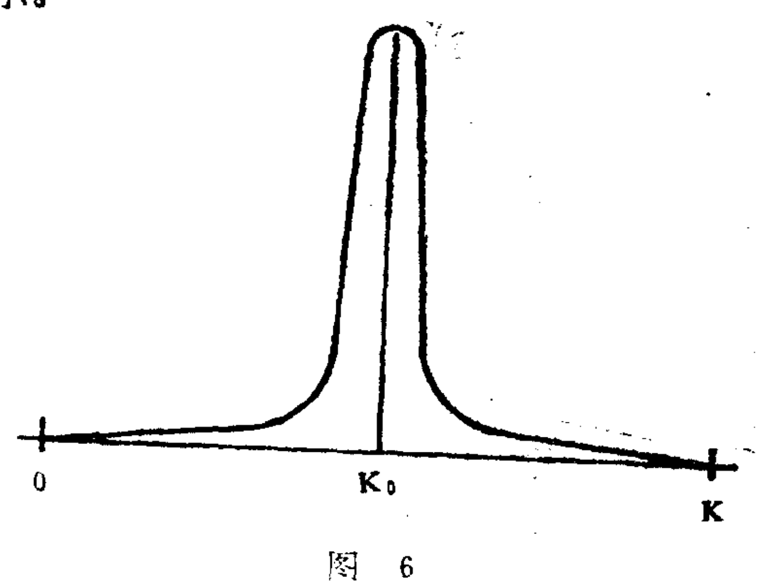


图 6

上述事实显然意味着，**如果进行一批n次试验，且n为大数，那么可以满有把握地预料，事件A出现次数k将非常接近于其最大概率对应值k_0。**

这就是十八世纪初确立的著名的伯努利定理。它是概率论最重要的定理之一。证明这个定理起初需要复杂的数学工具。十九世纪中期著名的俄国数学家 П·Л·车比雪夫首先找到了简捷的数学证明。下面就来研究车比雪夫的证明。

§17 伯努利定理求证

我们业已知道，当试验次数n相当大时，事件A出现的最大概率对应值k_0几乎与np的数值没有区别，其中p表示单个试验中事件A出现的概率。因此，只要能证明下面的情

况就行了：当试验次数相当大时，在绝大多数情况下，事件A出现的次数k与np的差别很小——不大于n的任意分之一（比如，不大于0.01n或0.001n，或者说，不大于 εn，其中，ε为任一小数）。换句话说，应该证明，对于足够大的数值n来说，下面的概率将是任意地小、

$$P(|k-np|>\varepsilon n) \tag{1}$$

欲求证上式 我们发现，根据加法定理，在数k距离np为 εn的范围内，式（1）的概率值等于所有k值（k与np只差 εn）相应的概率$P_n(k)$的和。在图7中，其概率之和等于距np左右两端为 εn的AB截线内所有垂直线段之和。因为所有垂直线段的总和为1（事件的全系统概率之和），则这意味着，在AB区间上，所有概率之和几乎等于1，而在AB区间以外，所有概率之和只占微不足道的一个小数。因此，

$$P(|k-np|>\varepsilon n)=\sum_{|k-np|>\varepsilon n} P_n(k) \tag{2}$$

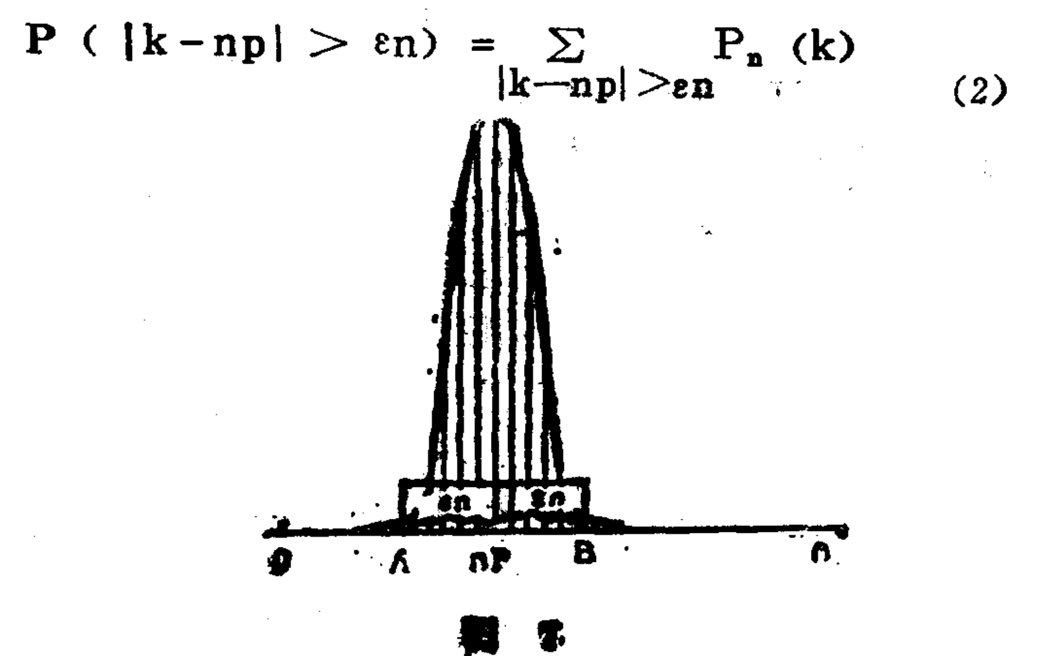


图 7

现在，再来研究车比雪夫推论。因为在所描述的概率之和中每一加数都有

$$\left|\frac{k-np}{\varepsilon n}\right|$$

即意味着

$$\left(\frac{k-np}{\varepsilon n}\right)^2>1$$

如果用$\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k)$来代替每个加数$P_n(k)$，就可以扩大这个和数。于是：

$$P(|k-np|>\varepsilon n)<$$

$$\sum_{|k-np|>\varepsilon n}\left(\frac{k-np}{\varepsilon n}\right)^2 P_n(k)$$

$$=\frac{1}{\varepsilon^2 n^2}\sum_{|k-np|>\varepsilon n}(k-np)^2 P_n(k)$$

显然，上式右边的和式还可以扩大。如果扩大k的范围，使其不仅从$np-\varepsilon n$到$np+\varepsilon n$，而是从o到n，这样，下列不等式更可成立：

$$P(|k-np|>\varepsilon n)<\frac{1}{\varepsilon^2 n^2}\sum_{k=0}^{n}(k-np)^2 P_n(k) \tag{3}$$

与以前各式不同的是，式（3）中的和可以准确计算；车比雪夫方法就在于将能作准确计算的和式取代了难以估算的和式。

现在着手计算，不管这个计算有多长，这只不过是技术方面的困难，凡懂代数的人就可以克服它。我们已经完全采用了车比雪夫的思想，这就是从等式（2）演变到了不等式（3）。

首先，容易得到：

$$\sum_{k=0}^{n}(k-np)^2P_n(k)=\sum_{k=0}^{n}k^2P_n(k)$$

$$-2np\sum_{k=0}^{n}kP_n(k)+n^2p^2\sum_{k=0}^{n}P_n(k) \qquad (4)$$

上式右边第三部分是全事件系统概率之和，其值为1，这样只需计算以下两式：

$$\sum_{k=0}^{n}kP_n(k) \text{ 和 } \sum_{k=0}^{n}k^2P_n(k)$$

上面两个式子里，对应于$k=0$的加数等于零，因此可以从$k=1$加起。

1．为了计算这两个和式，用第五章公式（5）来表达$p_n(k)$，得到

$$\sum_{k=1}^{n}kP_n(k)=\sum_{k=1}^{n}\frac{kn!}{k!\ (n-k)!}p^k(1-p)^{n-k}$$

因为 $n!=n(n-1)!$ 和 $k!=k(k-1)!$，则有

$$\sum_{k=1}^{n}kP_n(k)=np\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!}$$

$$\times p^{k-1}\ (1-p)^{(n-1)-k-1}$$

或者在上式右边，令$k-1=l$并当k从1变到n时，l从0变

到（n－1），上式可写为

$$\sum_{k=1}^{n} kP_n(k) = np\sum_{l=0}^{n-1}\frac{(n-1)!}{l!(n-1-l)!}p^l(1-p)^{n-1-l}$$

$$= np\sum_{l=0}^{n-1} P_{n-1}(l)$$

上式中$\sum_{l=0}^{n-1} P_{n-1}(l)$是一个全事件系统概率之和，表示在（n－1）次试验中，事件出现的全部可能数值。显然其和等于1。

这样一来，关于$\sum_{k=0}^{n} kP_n(k)$之和，我们得到了非常简单的表达式：

$$\sum_{k=0}^{n} kP_n(k) = np \qquad (5)$$

2．为计算第二个加式，首先求$\sum_{k=1}^{n} k(k-1)P_n(k)$的数值。因为当k＝1时，此式的和显然等于零，则可从k＝2开始加起。注意到$n! = n(n-1)(n-2)!$ 和$k! = k(k-1)(k-2)!$，和前面一样，令k－2＝m，从而得到：

$$\sum_{k=1}^{n} k(k-1)P_n(k)$$

$$= \sum_{k=2}^{n} k(k-1)P_n(k) = \sum_{k=2}^{n}\frac{k(k-1)n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

$$=n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)![(n-2)-(k-2)]!}\times p^{k-2}(1-p)^{(n-2)-(k-2)}$$

$$=n(n-1)p^2\sum_{m=0}^{n-2}\frac{(n-2)!}{m!(n-2-m)!}p^m(1-p)^{n-2-m}$$

$$=n(n-1)p^2\sum_{m=0}^{n-2}P_{n-2}(m)=n(n-1)p^2 \qquad (6)$$

这是因为$\sum_{m=0}^{n-2}P_{n-2}(m)=1$，表示n－2次试验中事件出现的全部可能数，即是全事件系统概率之和。

根据（5）、（6）两式，最后得到：

$$\sum_{k=1}^{n}k^2P_n(k)=\sum_{k=1}^{n}k(k-1)P_n(k)+\sum_{k=1}^{n}kP_n(k)$$
$$=n(n-1)p^2+np=n^2p^2+np(1-p) \qquad (7)$$

至此，需要计算的两个加式之和计算完毕。把（5）、（7）两式结果代入公式（4），最后得到：

$$\sum_{k=0}^{n}(k-np)^2P_n(k)=n^2p^2+np(1-p)$$
$$-2np\cdot np+n^2p^2=np(1-p)$$

把这个非常简单的式子代入不等式（3），得到

$$P(|k-np|>\varepsilon n)<\frac{np(1-p)}{\varepsilon^2n^2}=\frac{p(1-p)}{\varepsilon^2n} \qquad (8)$$

该不等式证明完毕。虽然数ε可选取任意小，不过，一旦选

定就不再更换，而试验次数n，根据我们的观点，可为任意大的数。因此，分式$\frac{p(1-p)}{\varepsilon^3 n}$可为任意小，这是因为，随着n值的增长，分父可为任意大，而此时，分子数值不变。

例如　设$p=0.75$，于是$1-p=0.25$，则

$$p(1-p)=00187<0.2$$

取$\varepsilon=0.01$，那么，代入不等式（8），有

$$P\left(\left|k-\frac{3}{4}n\right|>\frac{1}{100}n\right)<\frac{0.2}{0.0001n}=\frac{2000}{n}$$

如果$n=00000$则

$$P(|k-150000|>2000)<0.01$$

实际上，这意味着以下事实：如果某种生产在规定的工艺过程条件下，平均有75％的产品具有某种特性（例如，属于一级产品），那么，在200000件产品中，有148000到152000件产品具有该种特性，这种情况的概率为0.99（实际上几乎千真万确）。

对此须作两点说明：

1）不等式（8）对概率$P(|k-n|>\varepsilon n)$作了非常粗略的估计。事实上，特别是当n为大数时，这个概率值要小得多。因此，实际上要采用较为准确的估算方法。不过，这类方法的论证却是相当复杂的。

2）当概率p很小，或反之，或p接近于1时，用不等式（8）作出的估算是较为准确的。

如果刚才所举的例子里，产品具有某种特性的概率$p=0.95$，那么，

$$1-p=0.05 \qquad p(1-p)<0.05$$

因此，选取 $\varepsilon=0.005$、$n=200000$，求得

$$\frac{p(1-p)}{\varepsilon^2 n}<\frac{0.05\times1000000}{25\times200000}=0.01$$

这与前例结果相同。不过，此时，εn不等于2000，而仅为1000。因为$np=190000$所以得出：在总数为200000件的产品中具有这种特性的产品数在189000到191000件之间是确信无疑的。

这样一来，当$p=0.95$时，实际上从不等式(8)可知，对于预期数量的具有某种特性的产品而言，由于 $P(|k-190000|>1000)<0.01$，因此，其区间长度比 $p=0.75$ 时的小两倍。

习题 已知某工业部门有四分之一的工人受过中等教育。任选200000名工人进行某种调查。求1)在任选的200000名工人中，受中等教育的工人人数最大概率对应值；2) 这种工人的人数与最大概率对应数之差不大于1.6%的概率。

解题时，我们的出发点是：对任选的200000名工人中的每个工人而言，受过中等教育的概率等于四分之一（"任选"一词的含义就在于此）。这样一来，$n=200000$，$p=\frac{1}{4}$，

$k_0=np=50000$；$p(1-p)=\frac{3}{16}$。求 $|k-np|<0.016np$ 或 $|k-np|<800$的概率，式中，k为受中等教育工人数。

选择 ε时，应使 $\varepsilon n=800$。由此得出：

$\varepsilon=\frac{800}{n}=0.004$。由（8）式得到：

$$P(|k-50000|>800)$$

$$<\frac{3}{6\times 0.000016\times 200000}\approx 0.06$$

由此 $P(|k-50000|<800)>0.94$

答：所求最大概率对应值等于50000，所求概率大于0.94。事实上这个概率接近于1。

第二部分　随机变量

第七章 随机变量与分布规律

§18 随机变量的概念

前面各章里，已经多次碰到过，其数值不是恒定的，而是在各种随机影响下随之而变的量。例如，在每一百个新生儿中的男孩数量，不可能都是相同的。又如，某种棉花纤维长度不仅在不同的产地，而且在同一棉株上，其变化也是非常之大的。

现再举几个例子：

1． 尽管用同一武器，同一瞄准器，在不变的条件下进行射击，发现子弹却射中不同的位置，这种现象称为子弹"散布"。在事先无法估计的各种随机因素影响下，每次子弹射击的距离都不一样。

2． 气体分子运动速度是变化的，它随着与其他分子互相碰撞而改变，由于每个气体分子可能与其他分子相碰或不相碰，因此，分子速度的变化具有纯随机特性。

3． 每年落到地球表面的陨星数不是固定的，而是受一系列随机因素的影响，变化很大。

4． 某地区种植的小麦麦粒的重量不等于某一确定的量，而是变化的，鉴于决定小麦生长的各种因素的影响（如某种小麦生长的土壤质量，光照条件，水利情况等等）不可能都考虑到，因此，麦粒的重量是随机变化的量。

尽管上述各例均不一样，但从我们感兴趣的观点来看，

它们都属于同样一种情况。在每个例子里，所关心的是表征某种作业（例如：陨星的统计、纤维长度测量）结果测量值；在不同作业时，尽管力求使它们的条件都一样，但是，由于在这些作业的条件里，存在着被人们忽视的随机差异，因此，其中每个量值可为不同数值。在概率论里把这种量称为**随机变量**；上面列举的各例足以使我们相信，研究随机变量对于将概率论应用于各个不同领域里的实践中去是多么重要！

知道某个随机变量，并不意味着知道了该变量的数值。例如，知道电容器在击穿之前工作了5324小时，因此，电容器连续工作时间为一确定数值，并不再是一个随机变量了。要对随机变量有一个尽可能全面的概念，应该知道些什么呢？显然，首先应该知道它的一切数值。

例如，电灯泡的使用期，经试验最短为2306小时，最长为12108小时，那么，电灯泡的使用期就可以是这两个数之间的任一数值。例3也很明显，在一年内落到地球表面的陨星数可以是任意一个正整数，即0、1、2、3、……等。

但是仅仅知道随机变量的各种可能数值，却还不能对它进行实际的评价。例如在第二个例子里，如果在两个不同温度下研究气体，其分子运动速度的数值可能相同，但在这两种温度下得到的分子运动速度数据并不能作出任何比较性的评价。

不同温度的气体是有本质区别的。在两种温度下，气体的分子运动速度可能出现相同数据时，那么，以这样的分子运动速度来表征两种不同温度的气体的差别就不可能。如果要评价气体的温度特性，只知道分子某种可能的运动速度是

不够的，还必须要知道有多少分子以这样的速度运动，即必须知道分子以这种速度运动的概率。换句话说，必须知道我们感兴趣的随机变量的各种可能数据的概率。

§19 分布规律的概念

首先举个最简单的例子。对图8所示的靶子进行射击，射中Ⅰ区给3分，射中Ⅱ区给2分，射中Ⅲ区给1分㊀。

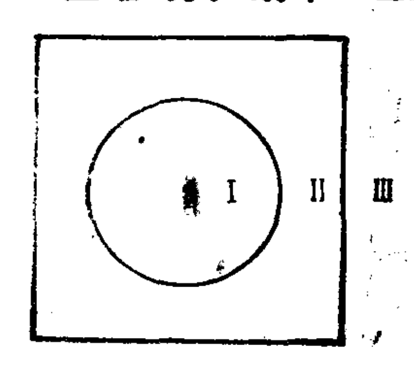


图 8

作为随机变量，我们研究单个射击的得分情况。这时，可能的得分数是1，2，3。用p_1，p_2，p_3相应地表示这三个数值的概率，比如，p_3表示命中靶子Ⅰ区的概率。可是，随机变量的可能数值对所有射手都是相同的，而对不同射手来说，概率p_1，p_2，p_3可能完全不相同。射击差别显然就由这个差别而定。例如，优秀射手命中概率可能为（比方说）$p_3=0.8$，$p_2=0.2$，$p_1=0$；中等射手，$p_3=0.3$，$p_2=0.5$，$p_1=0.2$；较差射手，$p_3=0.1$，$p_2=0.3$，$p_1=0.6$。

㊀ 读者可能提出异议，说：对射中Ⅲ区即脱靶的情况不应给1分。不过，如果有射击权即可得分的话，那么，射击最差的人本身就已得了一分。

如果在某次打靶中射击了12次，那么，从0到12的所有整数都可成为射中Ⅰ、Ⅱ、Ⅲ区每一区的可能数值；不过，这个事实本身并不能使我们判断打靶的质量。相反，如果除了可能的命中数值以外，还给出了这些数值的概率，也就是在12次为一组的射击中打中这个或那个区域的命中数。那么，对这次射击质量则可以形成一种清晰的概念。

显然，所有情况都是这样的；若知道随机变量各种可能数值的概率，同样就可以知道，果真要出现的较有利的或者较不利的随机变量数值；当然，这种数值就足以判断与该随机变量相关的某项作业的有效性和优质性。实际表明，知道了被研究的随机变量的一切可能数值的概率，就足以解决同评价该随机变量相关的任何问题，并可作为相应作业的优质性指标。于是可以得出结论，为了完整表征这种或那种随机变量，应该知道的必要和充分条件是：

1）该随机变量的一切可能数值；

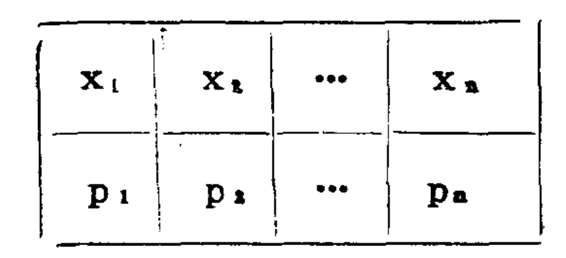
2）每个数值的概率。

由此可见，用两行表格的方式可方便地给出随机变量，上一行包含以某一顺序排列的随机变量的可能数值，下一行是与可能数值一一对应的概率。比如，上述例子里的优秀射手在一次打靶中所得的分数作为随机变量可列表如下

1	2	3
0	0.2	0.8

（Ⅰ）

在一般情况下，随机变量可能数值是x_1、x_2、……、x_n，而相应的概率是p_1、p_2、……、p_n，于是得出下表



x_1	x_2	…	x_n
p_1	p_2	…	p_n

列出这种表格，也就是建立随机变量的一切可能数值及其概率，这就意味着建立该随机变量的**分布规律**。知道某个随机变量的分布规律，就可以解决所有的与之相关的问题。

例题　在一次打靶中，某射手所得分数数值具有表（Ⅰ）所示的分布规律；另一个射手射击所得分数数值的分布规律如表（Ⅱ）

1	2	3
0.2	0.5	0.3

（Ⅱ）

求两射手射击得分之和的分布规律。

显然，这里所指的和是随机变量。我们的任务是建立它的分布规律表。为此，应该研究两个射手联合射击的所有可能结果。并把这些结果放置在下列表格之中，表中每个结果的概率都按独立事件乘法定律计算，x表示第一个射手射击得分数，y表示第二个射手射击得分数。

射击结果序号	x	y	x+y	结果的概率
(1)	1	1	2	0×0.2=0
(2)	1	2	3	0×0.5=0
(3)	1	3	4	0×0.3=0
(4)	2	1	3	0.2×0.2=0.04
(5)	2	2	4	0.2×0.5=0.1
(6)	2	3	5	0.2×0.3=0.06
(7)	3	1	4	0.8×0.2=0.16
(8)	3	2	5	0.8×0.5=0.4
(9)	3	3	6	0.8×0.3=0.24

上表说明，我们感兴趣的x+y之和可以为数值3，4，和6；而不可能为数值2，这是因为该数值的概率为零㊀。在(2)、(4)两种结果里，有x+y=3；要使x+y之和为3，就应该出现(2)或(4)两种结果中的一种，根据加法定理，这种概率等于这两种结果的概率之和，即等于0+0.04。要使x+y=4，就必须出现(3)、(5)或(7)三种结果中的一种。因此，这个等式出现的概率等于0.26（也是根据加法定理）。用同样的方法，求出x+y之和等于5的概率为0.06+0.4=0.46；而x+y之和为6的情况仅在结果(9)中出现，其概率为0.24。于是，对于随机变量x+y，得到了

㊀ 当然，可以认为2是x+y之和的一个可能数值，其概率为零，这就象表（Ⅰ）中对数值1所做的规定一样

下列可能的数值表：

3	4	5	6
0.04	0.26	0.46	0.24

（Ⅱ）

表（Ⅱ）完全解答了所提的问题。

表（Ⅱ）中所有四部分概率之和为1，当然，每个分布规律都应具有这种性质，这是因为此处指的是关于随机变量所有可能数值的概率之和，即是关于某个全事件系统的概率之和。分布规律的这个特性可作为检验计算结果正确性的一种简便验算方法。

第八章 均 值㊀

§20 随机变量均值的确定

上面已经谈到，两个射手联合射击，根据随机状况可能得到3分，或者4分，或者5分，或者6分。这四种可能结果的概率，如表（Ⅱ）所示。如果问："两个射手进行一次双人射击可得多少分呢？"由于每次射击有每次的结果，因此这个问题不能回答，不过，为了评价两个射手的射击质量，需要的不是单个射击的结果（这种结果可能是随机的），而是整批射击结果的**平均值**。这两个射手平均一次射击可得多少分呢？这个问题提得完全合理，并且对它可作出明确的回答。

我们这样求解：如果这对射手进行一百次双人射击，则如表（Ⅱ）所示，

在一百次射击中约有4次得3分

在一百次射击中约有26次得4分

在一百次射击中约有46次得5分

在一百次射击中约有24次得6分

于是，平均每一百次双人射击得分总数为

$$3\times4+4\times26+5\times46+6\times24=490$$

㊀ 也称为数学期望。——校者注

用100除这个总数，得到平均每次射击得4.9分。这就是本问题的答案。我们发现，不用100除得分总数（490），而在做加法前，用100除每一个被加数，其总和直接就是一次射击的平均得分数。这种计算非常简单，用100除所有加数中的第二个乘数即可。可是，这些乘数都是用表（Ⅱ）所列的概率乘以100而得到的。要将它们再除以100，于是又还原成这样一些概率。这样一来，一次射击得到的平均得分数可用下式表示：

$$3\times0.04+4\times0.26+5\times0.46+6\times0.24=4.9$$

显而易见，上一等式左边加法是根据表（Ⅱ）按很简单的规律建立起来的。这个规律是，用可能数值表中上一行的每一个数乘以下一行的与之对应的概率，再把所有乘积加起来即可。

现在作一般形式的研究。假设，某种随机变量如下表所示

x_1	x_2	…	X_k
p_1	p_2	…	p_k

我们假定：如果变量x中，数值x_1的概率等于p_1，则这意味着在一批由n次组成的作业中，数值x_1将大约出现n_1次，这里，$n_1/n=p_1$，由此，$n_1=np_1$；同理，数值x_2在这种情况下大约出现$n_2=np_2$次，……，数值x_k大约出现$n_k=np_k$次。

于是，一批几次作业中

当$x=x_1$时，平均包含$n_1=np_1$次作业，

当$x=x_2$时，平均包含$n_2=np_2$次作业，

……………………………………………

当$x=x_k$时，平均包含$n_k=np_k$次作业。

在全部n次作业中，变量x数值之和大约是

$$x_1n_1+x_2n_2+\cdots+x_kn_k=n\ (x_1p_1+x_2p_2+\cdots+x_kp_k)$$

将上式除以n即可得到与单个作业相应的随机变量的均值$\overline{x}$

$$\overline{x}=x_1p_1+x_2p_2+\cdots+x_kp_k$$

这样一来，得出了以下重要规则：**为了得到随机变量均值，应该把随机变量的所有可能值乘以对应的概率，然后，把所得的乘积相加。**

知道随机变量的均值，可以给我们带来什么好处呢？为确切地回答这个问题，先来研究几个例子。

例1 再来研究两个射手的射击问题。他们射击得分的数值是一种随机变量，其分布规律是，第一个射手为表（I）所示，第二个射手为表（I）所示（见第69、70页）在这两个表里，一个引人注目的问题是，第一个射手比第二个射手要射得好，第一个射手最优秀的结果（3分）的概率比第二个射手的大得多；与此相反，第二个射手较差射击结果的概率比第一个射手的要大。不过，这种比较并不能使我们满足，因为它是一种纯定性的描述。这种描述缺乏象温度直接评价物体加热程度那样的度量标准，不能用自己的量值直接评价这个或那个射手的射击水平。假如没有一种评价尺度，

我们就会经常处于靠直接研究得不出任何答案，或者得出一个可能有争议的答案的窘境。比如，如果用下面两个表代替表（Ⅰ）、表（Ⅱ）。

1	2	3
0.4	0.1	0.5

（Ⅰ′）

1	2	3
0.1	0.6	0.3

（Ⅱ′）

其中，表（Ⅰ′）是第一个射手的，表（Ⅱ′）是第二个射手的。不过，根据这两个表难于用一种方法确定那个射手射击比较优秀。当然，第一个射手优秀成绩（3分）的概率比第二个射手的要大。但与此同时，第一个射手最差成绩（1分）的概率也比第二个射手的大。相反，第二个射手成绩为2分的概率比第一个射手的要大。

现在，按前面指出的规则，建立每个射手得分数的均值：

1）第一个射手为

$$1\times0.4+2\times0.1+3\times0.5=2.1$$

2）第二个射手为

$$1\times0.1+2\times0.6+3\times0.3=2.2$$

由此看到，第二个射手平均得分比第一个射手的略高。实际上这意味着，在多次射击时，第二个射手的射击成绩比第一个射手的要好。现在可以满怀信心地说，第二个射手射击成绩较为优秀。这是射击得分数均值给了我们一个便利的尺度，用它，可以轻而易举地、毋庸置疑地比较各个射手的技艺。

例2 在组装精密仪器时，为了较准确地调整某个器件，

根据工作顺手与否，可以提出1、2、3、4或5个备件。这样一来，实现合格组装所必要的备件数x是随机变量，其可能值为1、2、3、4、5。这些值的概率如下表所示

1	2	3	4	5
0.07	0.16	0.55	0.21	0.01

我们面临的任务是向工人提出组装20台仪器㊀所必要的器件数量。要想大体地估算这个数，不能直接采用上表——因为该表格表明在不同的情况下它是不一样的。不过，如果求一个仪器必要的备件x数的均值 $\bar{x}$，并乘以20，显然就可大体得到所求数的数值。求得

$$\bar{x}=1\times0.07+2\times0.16+3\times0.55+4\times0.21+5\times0.01$$
$$=2.93$$

$$20\bar{x}=2.93\times20=58.6\approx59$$

如果器件的实际消耗超过期望数，为便于装配工人有为数不多的备件待用，实际上，给他60～65个器件即可。

在上述各个例子里，当需要对某一随机变量作某种大体估算时，我们就要了解事情的规律。但是，仅用这种表格不能作出上面的估算，表格只是说明随机变量可以采用某一数值及其对应的概率。不过，用该表算出的随机变量均值已经能够作出这样的评价，因为它正是在一批连续作业中随机变

㊀ 假定，组装一台仪器的报废器件不再组装其他仪器。

量平均采用的数值。我们发现，当涉及到成批的或许许多多的重复作业时，从实用角度而言，均值恰如其分地描述了随机变量的特性。

习题1 我们进行了一系列使某一事件A出现的概率都为p的试验，而且，各次试验的结果相互独立。求在一批n次试验中事件A出现次数的均值。

在一批n次试验中，事件A出现数是一个随机变量，其可能数值为0、1、2、…n，而且据前所知数值k的概率为

$$P_n(k)=\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}$$

因此，所求均值等于

$$\sum_{k=0}^{n}kP_n(k)$$

求证伯努利定理时（第61页），曾经计算过这个和式，并已知其和等于np。当时业已证实，在n次试验中，当n为大数时事件A出现的**最大概率**对应数接近于np；现在看到，在试验数n为任一数时，事件A出现的均值恰好等于np。

这样一来，在这种情况下，随机变量最大概率对应值与随机变量出现次数的均值重合。值得提醒的是，对于任何随机变量而言，是否都会出现这种重合呢？一般地说，随机变量最大概率对应值可能与均值相差甚远。例如，对于分布规律如下表所示的随机变量而言，最大概率对应值是0，而均值是2.5。

0	5	10
0.7	0.1	0.2

习题2 作这样一些独立试验，在每次试验中，某一事件A可能出现的概率是0.8。在事件A首次出现之前，试验一直进行。总试验次数不超过4次，求所进行试验的平均数。

根据题目的已知条件，不得不进行的试验次数可能等于1、2、3或4。我们应该计算这四个数值中每一个的概率。若只要求进行一次试验，事件A就应该在第一次试验时出现。这个概率为

$$p_1 = 0.8$$

若要进行二次试验，事件A就不会在第一次试验时出现，而应在第二次试验时出现。根据独立事件乘法定理，这个概率应为

$$p_2 = (1-0.8)\times 0.8 = 0.16$$

若要进行三次试验，事件A就不在头二次试验时出现，而在第三次试验时出现。因此

$$p_3 = (1-0.8)^2\times 0.8 = 0.032$$

最后，当事件A在前三次试验中均未出现时，就需要进行第四次试验，因此

$$p_4 = (1-0.8)^3 = 0.008$$

这样一来，要进行的试验数是个随机变量，由以下分布规律而定：

1	2	3	4
0.8	0.16	0.032	0.008

它的均值为

$$1\times0.8+2\times0.16+3\times0.032+4\times0.008=1.248$$

比如，要作100次类似的观察，则可以算出，这时大约要进行$1.248\times100\approx125$次试验。

实际上与本题类似的情况经常碰到。棉花纤维的强度试验就是一例。当用同一包标准长度的纤维试样试验时（每次试验试样不超过四个），如果在载荷P作用下纤维不断开，则把它们列为优质一类。

习题3 有一广场呈正方形。根据航空测量，其边长等于350m，航空摄影测量质量由以下数字决定：

误差为0m的概率是0.42

误差为±10m的概率是0.16㊀

误差为±20m的概率是0.08

误差为±30m的概率是0.05

求广场面积的均值。

根据航空摄影测量的随机性，广场的边长为一随机变量，其分布规律如表所示：

320	330	340	350	360	370	380
0.05	0.08	0.16	0.42	0.16	0.08	0.05

（1）

㊀ 此处可这样认为，误差为+10m和−10m的概率都是0.16。以下情况类同。

由此立即可求出该变量的均值。在这种情况下甚至不必采用计算规则。由于表中两边同一误差有同一概率。根据对称性，显然，正方形边长均值等于被观察的数值，即350米。较为详细地说，均值的表达式包含下述各项

$$(340+360)\times0.16=[(350+10)+(350-10)]\times0.16$$
$$=2\times350\times0.16$$
$$(330+370)\times0.08=2\times350\times0.08$$
$$(320+380)\times0.05=2\times350\times0.05$$

因此，得

$$350\times(0.42+2\times0.16+2\times0.08+2\times0.05)=350$$

根据这种对称情况，似乎可以认为，正方形面积的均值应等于$350^2=122500m^2$。实际上，这个结论只有在随机变量平方的均值等于该随机变量均值的平方时才能成立。在本题中，广场面积可能为以下数值：

320^2　330^2　340^2　350^2　360^2　370^2　380^2

上述数值中实际上是哪一个呢？这取决于表（Ⅰ）所列七种情况中那一种具备条件，使得这七个值的概率与表（Ⅰ）的概率相同。简言之，广场面积分布规律应为

320^2	330^2	340^2	350^2	360^2	370^2	380^2
0.05	0.08	0.16	0.42	0.16	0.08	0.05

因此，广场面积均值为

$$320^2\times0.05+330^2\times0.08+340^2\times0.16+350^2\times0.42$$
$$+360^2\times0.16+370^2\times0.08+380^2\times0.05$$

利用对称性可简化计算。鉴于类似的简化，可能经常需要，因为我们来研究一下如何进行。重新改写上述表达式并进行计算：

$$350^2\times0.42+(340^2+360^2)\times0.16+(330^2+370^2)\times0.08+(320^2+380^2)\times0.05$$

$$=350^2\times0.42+[(350-10)^2+(350+10)^2]\times0.16+[(350-20)^2+(350+20)^2]\times0.08+[(350-30)^2+(350+30)^2]\times0.05$$

$$=350^2[0.42+2\times0.16+2\times0.08+2\times0.05]+2\times10^2\times0.16+2\times20^2\times0.08+2\times30^2\times0.05$$

$$=350^2+2\times(16+32+45)=122686$$

根据这种方法，一切计算都可采用心算。

我们看到，正方形面积的均值略大于（实际上，在这种情况下感觉不到差别）边长均值的平方（即$350^2=122500$）。由此容易证明下述普遍规律：**任何随机变量平方的均值都大于其均值的平方**。设有一随机变量x，其任意分布规律如下表：

x_1	x_2	……	x_k
p_1	p_2	……	p_k

那么，数值x^2的分布规律将为

x_1^2	x_2^2	…	x_k^2
p_1	p_2	…	p_k

这两个变量的均值相应为

$$\overline{x}=x_1p_1+x_2p_2+\cdots+x_kp_k$$

$$\overline{x^2}=x_1^2p_1+x_2^2p_2+\cdots+x_k^2p_k$$

则有

$$\overline{x^2}-(\overline{x})^2=\overline{x^2}-2(\overline{x})^2+(\overline{x})^2$$

由于$p_1+p_2+\cdots+p_k=1$，则右边三部分分别呈以下形式：

$$\overline{x^2}=\sum_{i=1}^{k}x_i^2p_i$$

$$2(\overline{x})^2=2(\overline{x})(\overline{x})=2\overline{x}\sum_{i=1}^{k}x_ip_i=\sum_{i=1}^{k}2\overline{x}x_ip_i$$

$$(\overline{x})^2=(\overline{x})^2\sum_{i=1}^{k}p_i=\sum_{i=1}^{k}(\overline{x})^2p_i$$

因此，

$$\overline{x^2}-(\overline{x})^2=\sum_{i=1}^{k}[x_i^2-2\overline{x}x_i+(\overline{x})^2]p_i$$

$$=\sum_{i=1}^{k}(x_i-\overline{x})^2p_i$$

由于等式右边所有加数之和不为负数，则

$$\overline{x^2}-(\overline{x})^2\geqslant 0$$

由此证明了上述普遍规律。

第九章　和与积的均值

§21　和的均值理论

实践中常常遇到这样的情况：已知两个（或多个）随机变量的均值，求这两个（或多个）变量和的均值。例如，有两个工厂生产同一产品，已知每天第一个工厂平均生产120件产品，第二个工厂平均生产180件。根据这些数据，能否确定每天两个工厂共同期望的产品数的均值呢？或者说，为了研究这个问题，除了这些已知条件外，还应该知道一些什么数据呢？例如，要否完全知道它们的分布规律？

事实上，要计算所有随机变量和的均值，知道各个被加数的均值就足够了，并且表达形式非常简单：**和的均值始终等于被加数均值的和。**

这样一来，如果x和y是两个完全任意的随机变量，那么

$$\overline{x+y}=\overline{x}+\overline{y}$$

在前一例中，x为第一个工厂生产的产品数，y为第二个工厂生产的产品数，$\overline{x}=120$，$\overline{y}=180$，即可知

$$\overline{x+y}=\overline{x}+\overline{y}=300$$

为了用一般方式证明这一规律，假设，变量x、y分别服从以下分布规律：

x_1	x_2	…	x_k
p_1	p_2	…	p_k

（Ⅰ）

y_1	y_2	…	y_l
q_1	q_2	…	q_l

（Ⅱ）

那么，变量x_i+y_j）式中$1\leq i\leq k$和$1\leq j\leq l$）的所有可能的数值也就是变量$x+y$的值。用p_{ij}表示x_i+y_j数值的概率。该量值我们不知道。它是$x=x_i$，$y=y_j$双重事件的概率，即是x为x_i，y为y_j时的概率。如果可以假定这两个事件互为独立的话，根据乘法定理，显然有

$$p_{ij}=p_i q_j \tag{1}$$

不过绝对不能这样假设。

因此，一般地说，式（1）不成立。表（Ⅰ），表（Ⅱ）也反映不出p_{ij}的数值。

根据一般规则，变量$x+y$的均值等于该变量所有可能数值与相应概率乘积之和，即

$$\overline{x+y}=\sum_{i=1}^{k}\sum_{j=1}^{l}(x_i+y_j)p_{ij}=\sum_{i=1}^{k}x_i\left(\sum_{j=1}^{l}p_{ij}\right)+\sum_{j=1}^{l}y_j\left(\sum_{i=1}^{k}p_{ij}\right) \tag{2}$$

现在，较仔细地研究一下和式$\sum_{j=1}^{l} p_{ij}$。

它是$x=x_i$，$y=y_j$的一切可能事件的概率之和，其中，数i不变，而j可从1到l。因为事件$y=y_j$，当j为不同值时，互不相容，则根据加法定理，和式$\sum_{j=1}^{l} p_{ij}$是l个事件中〔其中$x=x_i$，$y=y_j$（$j=1, 2, \cdots, l$）〕某一事件出现的概率。

不过，当说到“$x=x_i$，$y=y_j$（$1\leqslant j\leqslant l$）个事件中的任一事件出现”时，可简单地定为“$x=x_i$事件出现”。诚然，1）如果事件（$x=x_i$，$y=y_j$）（j等于所有可能数）中有一个出现，则显然，$x=x_i$事件出现；2）反过来说，如果$x=x_i$事件出现，由于y必定采用下列可能数y_1，y_2，…，y_l中的某一数值，那么，事件（$x=x_i$，$y=y_j$）($1\leqslant j\leqslant l$）中的任意一个就应出现。这样一来，$\sum_{j=1}^{l} p_{ij}$就是事件（$x=x_i$，$y=y_j$）（$1\leqslant j\leqslant l$）中任一个出现的概率，它直接地等于$x=x_i$事件的概率，即

$$\sum_{j=1}^{l} p_{ij}=p_i$$

同理得到

$$\sum_{i=1}^{k} p_{ij}=q_j$$

把这两个式子代入式（2），得到

$$\overline{x+y}=\sum_{i=1}^{k}x_ip_i+\sum_{j=1}^{l}y_jq_j=\bar{x}+\bar{y}$$

前式由此得证。

上述理论和证明可自动地推广到有三个或多个被加数的情况。根据刚才已经证明了的结果，可以写出

$$\overline{x+y+z}=\overline{x+y}+\bar{z}=\bar{x}+\bar{y}+\bar{z}$$

并可依此类推。

例 某工厂有n台机床，从每台机床生产的产品中选取一件产品。如果已知第一台机床出废品的概率等于 p_1，第二台——p_2．第 n台——p_n，求废品的平均数。

分析一个产品时，其废品数是个随机变量，它只能是两个数之一：如果产品是废品，其值为1；如果是正品，其值为0。就第一台机床而言，这两值的概率相应地等于p_1和$1-p_1$，废品平均数等于

$$1\cdot p_1+0\cdot(1-p_1)=p_1$$

第二台机床生产的产品的废品平均数等于p_2，依此类推。废品总数应是第一台、第二台及其他各台机床生产的废品数之和。因此，根据刚刚确定的均值加法定理，废品的平均数等于

$$p_1+p_2+\cdots+p_n$$

问题解毕。

其中，如果产生废品的概率各台机床都一样（$p_1=p_2=\cdots=p_n=p$），则废品总数的均值等于np。这个结果我们已经在第61页〔公式（5）〕里得出过。把解题时采用的繁琐计算方法与不需作任何计算能得出同样结果的简单推

算法比较一下是颇为有趣的。后者不仅方法简单而且普遍适用。在以前繁琐算法里，我们曾经假设：每个产品的生产结果是相互独立的，事实上，只有在这个假设之下，这个方法才能适用。现在，没有这个假设也行。这是因为在新的推算法里，我们采用了均值的加法定理，该定理适用于任何随机变量，不受任何限制。这样一来，不论在各个机床及其制造的产品之间是否有某种相互关联，只要所有机床生产废品的概率p都为同一数值，那么在取出的n件产品中，其废品数值等于np。

§22 积的均值理论

在随机变量之和里解决了的均值问题也经常会在随机变量之积里碰到。设随机变量x和y仍相应地服从表（I）、表（II）所示的分布规律。那么，xy的乘积也是随机变量，x_iy_j（$1\leqslant i\leqslant k$，$1\leqslant j\leqslant l$）乘积是该随机变量的可能值，x_iy_j值的概率等于p_{ij}。我们的任务是寻求一种在一切情况下用诸因子的均值表达变量xy的均值$\overline{xy}$的定理。不过，在一般情况下，解决这个问题似乎是不可能的。一般地说，知道均值$\overline{x}$和$\overline{y}$，并不能简单地确定$\overline{xy}$的量值（即在同样一些$\overline{x}$与$\overline{y}$值的条件下，$\overline{xy}$可能为不同的数值），因此，不可能有用$\overline{x}$和$\overline{y}$表达$\overline{xy}$的任何一种普通公式。

不过有一种重要的情况，如果某个表达式能够成立，并且，所得到的关系式相当简单。那么，就好办了。

我们规定，如果事件$x=x_i$，$y=y_j$在任一个i和j值情况

下都不相关，也就是说，在两个随机变量中，如果有一个取某一定值，并不影响第二个随机变量的分布规律。那么，称随机变量x和y为**相互独立**的。如果根据以上确定的思路，变量x和y相互独立，则

$$p_{ij}=p_iq_j\ (i=1,\ 2,\ \cdots,\ k;\ j=1,\ 2,\ \cdots l)$$

根据独立事件的乘法定理，有

$$\overline{xy}=\sum_{i=1}^{k}\sum_{j=1}^{l}x_iy_jp_{ij}=\sum_{i=1}^{k}\sum_{j=1}^{l}x_iy_jp_iq_j$$

$$=\sum_{i=1}^{k}x_ip_i\sum_{j=1}^{l}y_jq_j=\overline{x}\cdot\overline{y}$$

对于相互独立的随机变量而言，乘积的均值等于诸因子均值的乘积。

与和的均值理论情况一样，上述关于两个随机变量之积的定理可以自动地推广到任意个因子乘积的情况。这时，只需这些因子是互为独立的，即对其中部分变量给出某一定值应不影响其余变量的分布规律。

例1　用航空摄影测量法测量矩形面积，并测得该矩形边长为72m和50m。测量误差的分布规律并不知道，不过已经知道在矩形的这一边或那一边，同一变量的误差有相同的概率。于是，根据对称性（参看第80页习题8就易于证实）可知，矩形边长的均值与测量得到的结果相重合。如果这两种测量结果可当作相互独立的随机变量，那么，根据上述推导的乘法定理，矩形面积的均值将等于其边长的均值的乘积，即$72\times50=3600\text{m}^2$。不过，有时候有理由假定边长测量是相互关联的。比方说，两边都用同样一些准确校正的仪

器，测量就是相互关联的。如果长边测量数值明显超过实际长度，那么我们自然有根据推测：一般地说，这些测量仪器似乎会测得偏大数值，由于长与宽这两个数不是相互独立的，因此，在测量宽度时，测量数值偏大的概率也会增加。在这类情况下，面积的均值不可能等于矩形边长均值的乘积，要计算其面积需要附加其他的条件。

例2　电流沿导体流过。导体的电阻值按随机情况变化，该电流强度也按随机情况变化。已知导体电阻的均值等于25Ω，电流强度均值等于6A。计算沿导体流动的电流的电动势E的均值。

根据欧姆定律

$$E = RI$$

式中R——导体电阻，I——电流强度。

因为根据假设，$\overline{R} = 25$，$\overline{I} = 6$，While R与I是相互独立的，则有

$$\overline{E} = \overline{RI} = 25 \times 6 = 150V$$

第十章 偏离与平均偏差

§23 均值在表示随机变量时的不充分性

前面已经多次看到，随机变量的均值使我们对该变量有了一个初步的概念，就解决实际问题而言，这个概念在很多情况下是足够的。例如，要比较两个射手在竞赛中射击技术的高低，知道他们射击得分数的均值就够了。要比较两个不同宇宙微粒计数系统的效能，知道每个计数系统允许漏计的微粒数的均值也完全够了等等。在这些情况下，用均值而不用复杂的分布规律来描述随机变量颇为方便。就好象我们遇到的不是随机变量而是有确定数值的量一样。

不过常常碰到另一种情况，即随机变量较为重要的特性无论怎样都不能用其均值来确定，这时就需要较详细地知道随机变量的分布规律了。

在研究测量误差分布规律时就会碰到这类典型情况。设x为误差量，即被测之量测得的数值与其真值之差。没有系误差时，测量误差值（用$\overline{x}$表示）等于零。假定就是在这种条件下进行测量。试问误差将怎样分布呢？这种或那种量值的误差是否经常碰到呢？如果仅知道$\overline{x}=0$，则无法解决上述问题。由于误差量的均值等于零，则我们只知道可能有正、负误差，并且它们出现的机会大体相同。但是不知道

最主要的事实：即大多数测量结果是否接近于被测之量真的值。可能存在两种情况：1）相当可靠地信赖每个测量结果；2）这些测量结果基本上偏离真值，且正负误差值都很大。

当两个操作者用同一误差均值$\overline{x}$进行测量时，可能得到不同精确度的测量结果。其中一个人得出的测量结果可能比另一个人得出的有较大的系统偏离。这就意味着该操作者的平均误差较大，因此他的测量结果比另一个人的测量结果偏离被测量值要大。尽管两个操作者测量误差均值完全相同，但出现上述情况是完全可能的。

再研究一个例子。我们考察两种小麦的收获量，由于降水量、施肥、太阳照晒等随机条件不同，每平方米小麦收成受到很大的影响，因而是个随机变量。假定在同样条件下，每种小麦的平均产量都相同——每平方米产240克。仅仅根据平均产量的数字是否能判断被考察的某种小麦的质量呢？显然不能。因为经济效益大的小麦品种是其收获量受气候与其他随机因素影响较小的那一种。换句话说，即是收获量"偏离"较小的那一种。因此，在考察这种或那种小麦的收获量时，收获量可能的波动量比平均收获量具有更大的现实意义。

§24 度量随机变量偏离的各种方法

上述例子和其他一些类似的例子令人信服地表明，在很多情况下，要描述随机变量的本质特征，只给出随机变量的

均值是完全不够的。在给出随机变量均值的条件下，这些本质特征仍是未知的。为了描述那些特征，或者备有该随机变量的完整分布表（实际上这几乎总是复杂而又不方便的），或者除了随机变量的均值外，再试图引入一两个数，使这组为数不多的数综合起来能描述所研究的随机变量的一些最本质的特征。下面研究一下如何实现这种可能性。

正如上例表明，在很多情况下，最重要的问题是随机变量实际值相对其均值的偏差究竟有多大，大部分偏差是密集地散布在均值的周围（即彼此靠近）呢？还是相反——大部分偏差相对于均值有很大的偏离（在这种情况下，它们中的某些数值甚至彼此相差很大）？

下面的简例可明确地显示出这种情况。有两组随机变量，其概率分布如下表所示。

−0.01	+0.01
0.5	0.5

（Ⅰ）

−100	+100
0.5	0.5

（Ⅱ）

由表可知，两随机变量的均值都是零，可是第一组随机变量的值非常接近于零（而且彼此也接近），相反，第二组随机变量的值与零相差甚大（而且彼此相差也很大）。对于第一组变量，知道其均值，就大体可知其实际可能值了。对于第二组变量，其均值与实际可能值相差很大，因此，其实际可能值就不得而知。于是我们就说，第二种情况可能的偏离值比第一种情况的要大得多。

这样以来，我们的任务就是求能用合理方式给出随机变

量偏离程度的数，即偏这个数只是大体上表示出了随机变量与其均值偏差的大小。随机变量x与其均值$\overline{x}$的差$x-\overline{x}$本身显然也是一个随机变量；这个偏差的绝对值（不考虑符号，只表示偏差大小）$|x-\overline{x}|$也将是个随机变量。我们希望有一个能大体表征这种随机偏差$|x-\overline{x}|$的数，并能指明这种偏差可能有多大。求解这一问题有许多方法，实际上最常用的有下列三种：

1. **平均偏差** 人们总是自然而然地把随机变量$|x-\overline{x}|$的均值$\overline{|x-\overline{x}|}$作为它的概略值，这种偏差的绝对值的均值叫做变量x的**平均偏差**。如果随机变量x如下表所示

x_1	x_2	…	x_k
p_1	p_2	…	p_k

则随机变量$|x-\overline{x}|$的表为

$\|(x_1-\overline{x})\|$	$\|(x_2-\overline{x})\|$	……	$\|x_k-\overline{x}\|$
p_1	p_2	……	p_k

其中，$\overline{x}=\sum_{i=1}^{k}x_ip_i$。对于随机变量x的平均偏差$M_x$，得出以下公式：

$$M_x = \overline{|x-\overline{x}|} = \sum_{i=1}^{k} |x_i - \overline{x}| \ p_i$$

式中，仍有 $\overline{x} = \sum_{i=1}^{k} x_i p_i$。就表（Ⅰ）和（Ⅱ）给定的数值而言，$\overline{x}=0$，于是相应地有

$$M_{x_{\text{Ⅰ}}} = 0.01 \quad 和 \quad M_{x_{\text{Ⅱ}}} = 100$$

不过这两个例子平淡无奇，因为在这两种情况下偏差的绝对值似乎只能取一个数值，这样一来就丧失了随机变量的特征。

再来计算第76页表（Ⅰ′）和（Ⅱ′）确定的随机变量的平均偏差。计算表明，这些变量的均值相应地等于2·1和2·2，即彼此非常接近。第一个变量的平均偏差等于

$$|1-2.1| \times 0.4 + |2-2.1| \times 0.1 + |3-2.1| \times 0.5 = 0.9$$

而第二个变量的平均偏差等于 $|1-2.2| \times 0.1 + |2-2.2| \times 0.6 + |3-2.2| \times 0.3 = 0.48$

第二个变量的平均偏差几乎是第一个的1/2。实际上这自然意味着：虽然两个射手射击平均得分大约相同（从这一点出发可认为他俩射击技能相同），但是与此同时，第二个射手射击较稳定，其射点的偏离比第一个射手的要小得多，而第一个射手在平均射击得分相同的条件下射击特性不稳定，射击优、劣结果相差很大。

2．**均方差**　用平均偏差法度量偏差的概略值是很自然的事，与此同时，由于计算与评估带绝对值的数经常很复杂，有时甚至全然办不到，因此它实际上是一种很不方便的办法。所以求偏差量时通常采用其他方法——均方差法。

象随机变量x与其均值$\overline{x}$的偏差$x-\overline{x}$是随机变量一样，其偏差的平方$(x-\overline{x})^2$也是随机变量，仍采用前面的符号列出下表。

$(x_1-\overline{x})^2$	$(x_2-\overline{x})^2$	……	$(x_k-\overline{x})^2$
p_1	p_2	……	p_k

因此其平方的均值等于

$$\sum_{i=1}^{k}(x_i-\overline{x})^2p_i$$

根据这个数值可知$x-\overline{x}$差的平方大约等于什么数值，因此，对这个数值开平方，有

$$Q_x=\sqrt{\sum_{i=1}^{k}(x_i-\overline{x})^2p_i}$$

由此得到了能表示偏差程度的数值。Q_x称为随机变量x的均方差，它的平方即Q^2_x称为随机变量x的方差。当然，计算偏差值的这种方法比前面引进的平均偏差法较为巧妙，即采用迂回的办法首先求偏差平方的概略值，然后用开平方的方法再来求偏差值。但是，正如下一节要看到的，用均方差法Q_x能大大简化计算。正因为如此，统计员们在实践中大都采用均方差法。

实例 对于第76页表（Ⅰ′）和表（Ⅱ′）确定的随机变量。相应地有

$$Q^2_{x_I}{}' = (1-2.1)^2\times0.4+(2-2.1)^2\times0.1$$
$$+(3-2.1)^2\times0.5=0.89$$
$$Q^2_{x_{II}}{}' = (1-2.2)^2\times0.1+(2-2.2)^2\times0.6$$
$$+(3-2.2)^2\times0.3=0.36$$

因此

$$Q_{x_I}{}' = \sqrt{0.89}\approx0.94,\ Q_{x_{II}}{}' = 0.6$$

前面已经求出了这些变量的平均偏差

$$M_{x_I}{}'=0.9,\ M_{x_{II}}{}'=0.48$$

我们发现，如同平均偏差一样，第一个变量的均方差比第二个变量的要大得多。用平均偏差法和用均方差法度量偏离程度都能得出同一结论，即在两个变量中，第一个的偏离比第二个的要大。在这种情况下将采用哪一种方法呢？

这两种情况的均方差都比平均偏差大，由此可以推测，任何随机变量都是如此。事实上，随机变量 $|x-\overline{x}|$ 平方的均值Q_x，根据第85页已证明的公式，不可能小于随机变量 $|x-\overline{x}|$ 均值M_x的平方，即$Q^2_x\geqslant M^2_x$，由此得出$Q_x\geqslant M_x$。

3．中位差（概差）为确定偏离的量，经常（特别是在军事上）需要另外的方法，现用炮兵例子进行阐述。

假设炮兵从O点向OX方向（图9）进行射击。弹着点距发射点的距离x是随机变量，其均值表示"命中中心"C的位置（$OC=\overline{x}$），各个单发炮弹落点程度不同地散布在C点附近。

随机变量（炮弹飞行距离）与其均值的差值$x-\overline{x}$同时也是炮弹落在离命中中心C的差值，因此对 $|x-\overline{x}|$ 变量的所有

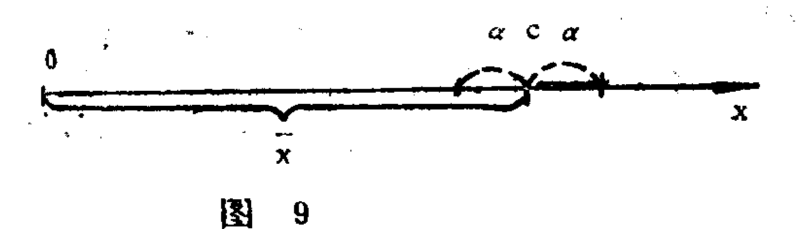


图 9

估算同时也就估算了炮弹在C点附近散布即偏离程度，于是，$|x-\bar{x}|$变量便成为射击质量的一个最重要的指标。

如果从C点向左、向右截取长度为α的很小一段距离，那么只有为数不多的炮弹落在距中点为C，长度为2α的距离之内，换句话说，$|x-\bar{x}|<\alpha$，当α很小时其概率也是很小的。但是现在把截取的距离扩张，即增大α的数值（须知它是一个任意确定的量）。截取的距离越大，落入其间的炮弹数将越多，因此单发炮弹落入该区间的概率也将越大。当α很大时，实际上所有炮弹都将落在这种区间之内，这样一来，随着α数值的逐渐增大，不等式$|x-\bar{x}|<\alpha$的概率从零增加到1。首先，当α很小时，很可能有$|x-\bar{x}|>\alpha$，即炮弹落在截取距离之外。然后，当α增大时，将有$|x-\bar{x}|<\alpha$，即炮弹落在截取距离之内。因此，当α数从很小的数值变到较大数值的某一处，该α数应该有这样一个值α_0，即在截取距离中心为C、长度为$2\alpha_0$的区间内、外，炮弹命中的概率相同。换言之，不等式$|x-\bar{x}|<\alpha_0$和$|x-\bar{x}|>\alpha_0$具有相同的概率，即两者的概率都为1/2(如果把等式$|x-\bar{x}|=\alpha$极小概率忽略不计的话）。当$\alpha<\alpha_0$时，两个不等式中第二个的概率较大；当$\alpha>\alpha_0$时，第一个的概率较大。因此，存在着一个唯一确定的数α_0，偏差的绝对值可能大于α_0或小

于α_0的概率相同。

α_0数值为多大——它取决于射击武器的质量，不难看到，与平均偏差、均方差相似，α_0数值可作为炮弹偏离的量度。实际上，如果α_0很小，则这意味着大炮发射的全部炮弹有一半落在C点附近的一段很小的截距之内，从而表明偏离较小。反之，如果α_0较大，则在C点附近截距较大，我们就应该认为，有半数炮弹将落在该截距的区间内，显然，这表明炮弹在中心附近的偏离较大。

人们通常把α_0数值称为变量x的**中位差或概差**。如果对于某一数值，$x-\bar{x}$之差的绝对值大于它也好，小于它也好，其出现的概率相等，那么，称这个数值为随机变量x的中位差或概差。虽然变量x的中位差（我们用E_x表示）在数值计算方面不比平均偏差M_x方便，也明显不如均方差Q_x方便，但是，炮兵为了评估所有的偏差，仍采用中位差E_x。下面将明白为什么这个方法在实际上易于推行。

§25 均方差理论

可以相信，用均方差法表征偏差量比其他所有方法——平均偏差法、中位差（概差）法等等具有更为可取的一些特性。下面介绍该方法的基本知识。

设有随机变量x_1，x_2，…，x_n，其均方差为q_1，q_2，…q_n。假定$x_1+x_2+\cdots+x_n=X$，若q_1，q_2，…q_n已知，且随机变量x_i（$1\leqslant i\leqslant n$）相互独立，求变量X的均方差Q。

根据均值的加法定理，有

$$\overline{X}=\overline{x_1}+\overline{x_2}+\cdots+\overline{x_n}$$

因此

$$X-\overline{X}=(x_1-\overline{x_1})+(x_2+\overline{x_2})+\cdots+(x_n-\overline{x_n})$$

由此

$$(X-\overline{X})^2=\left[\sum_{i=1}^{n}(x_i-\overline{x_i})\right]^2$$

$$=\sum_{i=1}^{n}(x_i-\overline{x_i})^2+\sum_{i=1}^{n}\sum_{\substack{k=1\\ i\neq k}}^{n}(x_i-\overline{x_i})(x_k-\overline{x_k}) \quad (1)$$

因为

$$\overline{(X-\overline{X})^2}=Q^2, \quad \overline{(x_i-\overline{x_i})^2}=q_i^2$$
$$(1\leqslant i\leqslant n)$$

利用均值的加法定理从而得到

$$Q^2=\sum_{i=1}^{n}q_i^2+\sum_{i=1}^{n}\sum_{\substack{k=1\\ i\neq k}}^{n}\overline{(x_i-\overline{x_i})(x_k-\overline{x_k})} \quad (2)$$

不过，因为变量x和x_k满足前面的假设，当$i\neq k$时互为独立，那么根据相互独立的变量均值的乘法定理，当$i\neq k$时

$$\overline{(x_i-\overline{x_i})(x_k-\overline{x_k})}=\overline{(x_i-\overline{x_i})}\ \overline{(x_k-\overline{x_k})}$$

等式右边两个因子的积等于零，其原因是

$$\overline{(x_i-\overline{x_i})}=\overline{x_i}-\overline{x_i}=0$$

这样一来，在等式（2）的第二个和式里，每个单独的和均为零，因此得到

$$Q^2=\sum_{i=1}^{n} q_i^2$$

即**相互独立的随机变量的和的方差等于它们方差之和。**

这样，在随机变量互为独立的情况下，除了均值的加法定理外，又增添了一个很重要的**方差的加法定理**。由此，得到均方差为

$$Q=\sqrt{\sum_{i=1}^{n} q_i^2} \qquad (3)$$

可以用各个相互独立的被加数的均方差简单地表示其和的均方差，这是均方差法优于平均偏差法、中位差法及其他偏差法的一个最重要的优点。

例1 如果某企业里每件产品可能为废品的概率为p，那么在n件产品中废品的平均数等于np。为了大体确定废品实际数与其平均数的偏差可能有多大，需求出废品数与np的均方差。最简便的方法是采用公式（3）。

废品总数的确可以看作制造每一产品时出现的废品数之和（就象在第78页研究类似例子时已经见过的一样）。若认为这些数是互为独立的随机变量，则根据方差的加法定理，要计算废品总数的均方差Q，可以采用公式（3），其中，q_1，q_2，…q_n表示在制造每一产品时出现的废品数的各个均方差。不过在制造第i个产品时废品数由下表确定。

1	0
p	1 − p

因此$\overline{x_i}=p$，并且

$$q_i^2=\overline{(x_i-\overline{x_i})^2}=(1-p)^2p+p^2(1-p)$$
$$=p(1-p)$$

于是

$$Q=\sqrt{\sum_{i=1}^{n}q_i^2}=\sqrt{np(1-p)}$$

问题由此得解。

将废品平均数np与其均方差$\sqrt{np(1-p)}$比较发现，当n值很大时，后一数值大大地小于前一数值，且只占平均数的一个不大的分数。例如当$n=60000$，$p=0.04$时，废品平均数等于2400，而均方差$Q=\sqrt{60000\times0.04\times0.96}=48$，于是，废品实际数大约偏离废品平均数2%。

例2 设想装配一种由n个零件组成的机械，这些零件沿某一轴一个紧靠一个排列，并且两端被一个零件框住（图10）。

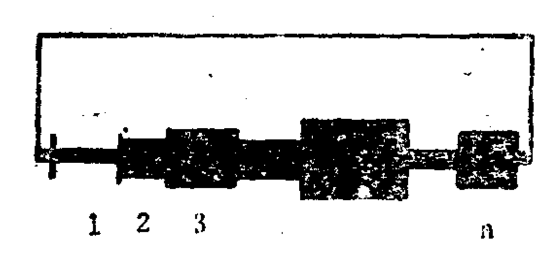


图 10

每个零件的长度可能与相应的标准长度有些差别，因此是个随机变量。假设这些随机变量是独立的。如果各零件的平均长度及其长度的均方差相应等于

$a_1, a_2, \cdots, a_n$ 和 $q_1, q_2, \cdots, q_n$

则n个零件组成的链的长度的均值和均方差为

$$a=\sum_{k=1}^{n} a_k \quad \text{和} \quad q=\sqrt{\sum_{k=1}^{n} q_k^2}$$

如果 $n=9$，$a_1=a_2=\cdots=a_9=10\text{cm}$。$q_1=q_2=\cdots=q_9=0.2\text{cm}$，则 $a=90\text{cm}$，

$q=\sqrt{9\times(0.2)^2}=0.6\text{cm}$。

于是我们发现，如果每个零件平均长度偏离其均值2%那么由这些零件组成的链的长度偏离其平均长度约为 $\frac{2}{3}\%$。当一些随机变量相加时其相对误差减少——这种情况在精密机械装配时起很大作用。诚然，如果每个零件的尺寸与给定的名义尺寸的偏差不互相补偿，那么装配机械时经常会碰到下列情况：外框零件长度小于被框住的零件链长，或者与此相反，外框零件留下特别大的间隙。在这两种情况下显然出现废品。采用减少“公差”的办法，即用减少实际尺寸与标定尺寸间的允许偏差的办法来与废品作斗争是不明智的，这是因为只要稍为提高加工精度，就会大大提高产品成本㊀。

㊀ 从技术上考虑得出的结论是：必须建立以概率论为基础的公差理论。

例3 假定在不变的条件下对某一随机变量进行n次测量。由于仪器状况、观察条件、空气状态的变化情况、防尘设备设置情况等等诸因素的影响，一般地说每次测量将得到不同的结果——产生随机测量误差。用x_1、x_2、……、x_n表示测量结果，作为标识符号，在每个x下注出测量序号。所有这些随机变量的均值都为$\overline{x}$，因为各次测量是在不变的条件下进行，显然应该假定各次测量的均方差q也是同样的。最后，同样认为变量x_1、x_2、…x_n是相互独立的。

现在求n次测量结果的算术平均值。

$$\xi=\frac{x_1+x_2+\cdots+x_n}{n}$$

这是一个随机变量，现求它的均值和均方差。

根据加法定理：

$$\overline{\xi}=\frac{1}{n}\ \overline{(x_1+x_2+\cdots+x_n)}$$

$$=\frac{1}{n}\ (\overline{x_1}+\overline{x_2}+\cdots+\overline{x_n})$$

$$=\frac{1}{n}\ (n\overline{x})\ =\overline{x}$$

这就是均值，其实它与各次单独测量所得到的完全相同。

其次，$x_1+x_2+\cdots+x_n$和的均方差根据方差的方法定理（3）有

$$Q=\sqrt{nq^2}=q\sqrt{n}$$

而随机变量ξ的均方差是上述和值的1/n，它等于

$$Q/n=q/\sqrt{n}$$

由此得出一个很重要的结果：**n个相互独立的同样分布的随机变量的算术平均值具有以下两个特点：**

1） 其均值与每个随机变量的均值完全相同。

2） 其均方差是每个随机变量均方差的$1/\sqrt{n}$。

比如，若被测量的变量的均值等于200m，而均方差等于5m，那么一百次测量结果的算术平均值ξ也为200m，但其均方差是单独测量的$1/\sqrt{100}=1/10$，即为0.5m。这样一来有根据期望，一百次实际测量结果的算术平均值比任一单独测量结果的算术平均值更接近于均值200m。由此可知：**大量相互独立随机变量的算术平均值的偏离比其中每个随机变量的要小很多倍。**

第十一章　大数定理

§26　车比雪夫不等式

前面已多次介绍过，知道随机变量的某种平均偏差（例如，它的方差），就可大致知道该随机变量的实际值偏离期望值有多大。至于偏差大的概率是多少，它并不能给出任何定量的评价，甚至近似的计算都不可能。有鉴于此，车比雪夫首先对此进行了简单的推算。从第96页随机变量x的方差表达式出发

$$Q_x^2 = \sum_{i=1}^{k}(x_i - \overline{x})^2 p_i$$

设α为任一正数，如果在上面的和式里，去掉 $|x_i - \overline{x}| \leqslant \alpha$ 的部分，只留着 $|x_i - x| > \alpha$ 的部分，则和必定减小，即

$$Q_x^2 \geqslant \sum_{|x_i - \overline{x}| > \alpha} (x_i - \overline{x})^2 p_i$$

如果用比 $(x_i - \overline{x})^2$ 小的数 α^2 代替每个被加数中的 $(x_i - \overline{x})^2$，则这个和要要减小

$$Q_x^2 \geqslant \alpha^2 \sum_{|x_i - \bar{x}| > \alpha} p_i$$

上式右边的和是在 $|x_i - \bar{x}| > \alpha$ 条件下，随机变量x的所有数值 x_i 的概率之和。根据加法定理，这是随机变量x得到其中某种数值的概率。换言之，这是实际所得偏差比α大时的概率 $P(|x - \bar{x}| > \alpha)$，上式可改写成

$$P(|x - \bar{x}| > \alpha) \leqslant \frac{Q_x^2}{\alpha^2} \quad (1)$$

这就是车比雪夫不等式。如果仅仅知道均方差 Q_x，仍旧可用上式评估偏差大于任意给定值α下的概率。诚然，车比雪夫不等式给出的评估常常很粗略，不过，该不等式除了理论上有极大的意义外，在实践中有时还被采用。

在上一节末尾曾经研究过这样的例子：测量结果的均值为200m，均方差为5m，在这些条件下实际得到偏差大于3 m的概率是很可观的（可以想象，它大于0.5；当然，其精确值只有在测量结果的分布规律完全知晓的情况下才能算出）。不过已经知道，对于一百次测量结果的算术平均值ξ来说，均方差为0.5m。因此由（1）式有

$$P(|\xi - 200| > 3) \leqslant \frac{(0.5)^2}{3^2} = \frac{1}{36} \approx 0.03$$

于是，对一百次测量的算术平均值而言，所得偏差大于3 m的概率已经很小了（事实上概率值还要小得多，以至可以全然不考虑这种偏差存在）。

在第101～102页例1里，检验60000件产品时得到废

品数的均值为2400件，均方差为48。对于实际废品数在2300到2500件之间（即$|m-2400|\leqslant 100$）的概率而言，车比雪夫不等式给出

$$P\{|m-2400|\leqslant 100\}$$
$$=1-P\{|m-2400|>100\}\geqslant 1-\frac{48^2}{100^2}\approx 0.77$$

实际上这个概率相当大。

§27　大 数 定 理

设有n个互为独立的随机变量x_1，x_2，…，x_n，它们具有同样的均值α和同样的均方差q。对于这些变量的算术平均值

$$\xi=\frac{x_1+x_2+\cdots+x_n}{n}$$

正如我们在第104页所见，均值等于a，均方差等于$q/\sqrt{n}$，因此在任一正数α下车比雪夫不等式给出

$$P(|\xi-a|>\alpha)\leqslant\frac{q^2}{\alpha^2 n}\qquad(2)$$

假设研究的问题是关于某一变量n次测量的算术平均值，并如§25末尾的例子所知，假定q＝5m，a＝200m，则有

$$P(|\xi-200|>\alpha)\leqslant\frac{25}{\alpha^2 n}$$

可以选α为很小的值，例如α＝0.5m，则

$$P(|\xi-200|>0.5)\leqslant\frac{100}{n}$$

如果测量次数n很大，那么上述不等式右边的数可以任意地小，比如，当$n=10000$时它等于0.01，于是，对于10000次测量的算术平均值，有

$$P(|\xi-200|>0.5)\leqslant 0.01$$

如果不计概率极小的事件，则可以说，在10000次测量中其算术平均值与200m之差的绝对值将不大于50cm。如果想更加接近，如10cm，则应该假定$\alpha=0.1m$，于是

$$P(|\xi-200|>0.1)\leqslant\frac{25}{0.01n}=\frac{2500}{n}$$

要使上面不等式右边的值小于0.01，不能把测量数定为10000（该数值现在不够），而应为250000。一般而言，不管α值多么小，只要取n足够大都可以使不等式（2）右边的数值达到预期的大小。这样一来，当n足够大时可以认为逆向不等式$|\xi-a|\leqslant\alpha$是何等接近于实际情况。

如果随机变量x_1，x_2，……，x_n相互独立，并且它们都有同样的均值a和同样的均方差，那么当n足够大时，变量$\xi=\frac{x_1+x_2+\cdots+x_n}{n}$与a之差值可以随意地小，它具有的概率可随意地接近于1（实际是足够地接近1）。

这是概率论基础理论中一个最简单的特殊情况，它称做**大数定理**，是由俄国著名数学家车比雪夫在上世纪中叶确立的。这个著名定理的深刻含意是：虽然单个随机变量的值经常（正如我们所知）比其均值小得多（具有相当大的偏离），但是大量随机变量的算术平均值却全然不同，这个变量的偏离很小，它的值非常接近于均值，且这种情况的概率极大。当然，这是因为在取算术平均值时，偏离两边的各随机

偏差相互抵消，因此在大多数情况下偏差之和就很小了。

实践中特别重视运用车比雪夫理论，因为选择少量材料就可判断大量同类材料的质量优劣。比如从一捆棉花的不同部位随机地抽取一小束（抽样），并用它来判断这捆棉花的质量。或者从一批谷物的不同位置随机地灌一小容器谷物，并用它来判断大批量谷物的质量㊀。根据这种选择来判断产品质量具有较大的准确性，因为尽管取出的谷物数量比储藏的全部谷物数小得多，但是就其本身而言仍是大数，根据大数定理就足以准确地评价某种谷物的平均重量，从而就可知该批谷物的质量。又根据一小束包含数以百计纤维而重量只不过十分之几克重的棉花，也可以准确地判断20普特（合327.6kg）一捆的棉花质量。

§28 大数定理的求证

到目前为止只研究了所有变量x_1，x_2，…具有同一均值及同一均方差的情况。但是，大数定理在更为一般的情况下也适用。现在研究以下这种情况。变量x_1，x_2，…的均值可能为任意数值（相应地用a_1，a_2，…表示），一般地说，它们各不相同。这时，变量

$$\xi=\frac{1}{n}(x_1+x_2+\cdots+x_n)$$

的均值显然将是变量

㊀ 譬如，灌入容器的谷物重量为100～200克重，而全部谷物为几十吨甚至几百吨。

$$A=\frac{1}{n}(a_1+a_2+\cdots+a_n)$$

并且根据车比雪夫不等式（1），当α为任一正数时，有

$$P(|\xi-A|>\alpha)\leqslant\frac{Q_\xi^2}{\alpha^2} \tag{8}$$

由此可见，一切都归结为评估随机变量Q_ξ^2，但是此处评估随机变量Q_ξ^2几乎和早先研究特殊情况时一样地简单。

Q_ξ^2是随机变量ξ的方差，ξ等于n个相互独立的随机变量之和除以n（当然仍然保存变量相互独立的假设）。因此根据方差加法定理有

$$Q_\xi^2=\frac{1}{n^2}\left(q_1^2+q_2^2+\cdots+q_n^2\right)$$

式中q_1，q_2，…相应地为随机变量x_1，x_2，…的均方差。现在假定这些均方差一般也是互不相同的，但是还要假定，无论随机变量怎么多（即无论数n有多大），所有这些变量的均方差总小于某个正数b。实际上，由于是在某种程度上同一类的变量相加，对不同的变量而言，其偏离程度不太悬殊，因此前面的条件总是能够满足的。

因而假定$q_i<b$（$i=1,2,\cdots$）。这时就有

$$Q_\xi^2<\frac{1}{n^2}nb^2=\frac{b^2}{n}$$

由此，根据不等式（8）有

$$(|\xi - A| > \alpha) < \frac{b^2}{n\alpha^2}$$

不管α有多么小，当所取的随机变量数n为足够大时，上述不等式右边的数值可达到随意小。显然这就证明了在一般情况下的大数定理。

总之，**如果变量x_1，x_2，……相互独立，而且它们的均方差都小于同样一个正数，则当n足够大时，其算术平均值**

$$\xi = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

偏离的绝对值可以为随意小的数，且这种概率也随意地接近于1。

这就是车比雪夫提出的一般形式的大数定理。

现在进而讨论一个重要的情况。假设测量某个变量a。在同样一些条件下重复测量时，测量者得到了一些不完全重复的数值x_1，x_2，…，x_n。取算术平均值作为a的近似值，即

$$a \approx \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

试问，如果试验的次数足够多，是否可以得到随意准确的数值a呢？

如果测量时没有系统误差，即如果

$$\overline{x}_k = a \text{（当}k = 1, 2, \cdots n\text{，时）}$$

而且这些数值本身不存在不准确性，即是说测量时在仪器上读出的数值就是实际上的数值。那么，上述问题的回答将是肯定的。如果仪器不能给出准确的读数而大于某个数值δ，比方说刻度尺上刻度的宽等于δ，那么读数精度就不可能高

于±δ。在这种情况下算术平均值显然与每个x_k一样，将具有同样的不准确度δ。

这个说明告诉我们，如果仪器给出的测量结果具有某一不准确度δ，那么试图用大数定理的方法得到准确度高的数值a。将是一种错误的想法，而且在这种情况下所作的计算本身将是一种空洞的算术游戏

第十三章　正态分布规律

§29　问题的提出

大量的自然现象、生产和操作过程实质上都是在这些或那些随机变量的参与下发生的。在自然现象、生产和操作过程尚未完成之前常常可能知道这些随机变量的分布规律，即知道随机变量的可能数值及其概率。如果随机变量有无数多个不同数值（炮弹飞行距离、测量误差等等），那么最好指出一整段数值对应的概率，而不必指出各单个数值对应的概率（例如，测量误差在－1mm到＋1mm、0.1mm到0.25mm等区间内对应的概率）。

这种看法并不改变问题的实质，即若要做随机变量的主人，要在力所能及的范围内掌握随机变量，就必须尽可能准确地知道它的分布规律。

要想知道所碰到的随机变量的分布规律，如果不作一般特性的一切推理和预测，如果在无任何预先假设的情况下处理每个随机变量，想用纯试验的方法求出随机变量固有分布规律的所有特征，这实际上是给自己提出了一项困难很大、几乎不能完成的任务。在每个新情况里，即使能确立新的未知的分布规律的一些最重要特点，也要进行大量的试验。因此学者们早就想寻求分布规律的一般模式，有了它就可以明智地对日常遇到的大量随机事件（纵然不是全部）进行预见、期望和推测。在理论上早已建立了这样一些模式，

并且已经得到了试验的证实。显然，根据理论研究成果及以前的所有试验，能够预见到新遇到的随机变量的分布规律属于何种模式该多方便啊！如果该预见正确，那么通常只需要少量的试验和观察就可以确定分布规律的全部特点了。

理论研究表明，在实际碰到的大量随机情况里，有充足依据预期到一种确定的分布规律。这 个 规 律叫**正态分布规律**。由于问题的复杂性，暂不考虑全部的证明及 准 确 的公式，本章只对该定理作些简要论述。

实际上在我们碰到的各种随机变量中很多都具有随机误差特性，或者至少易于归结为“误差”特性。例如，假设研究从某种火炮射出的炮弹的射程x。当然先假定有一个瞄准仪定标用的某一平均距离x。将瞄准仪定标在x_0上。$x-x_0$之差就是射程的“误差”，因此研究随机变量x就完全、直接地归结为研究$x-x_0$的“随机误差”。

不过这种每次射击都在变化的误差量取决于很多互不关联的因素。例如炮筒的随机振动，炮弹重量与外形不可避免的（尽管不大的）不一致性，引起空气阻力变化的大气条件的随机变化，瞄准中的随机误差（如果在每次单发射击前或每次连发射击前都重新瞄准）。所有这些因素及其他因素都能引起射程的误差。所有的局部误差都是相互独立的随机变量，而且每一随机变量的影响只占全部影响的很小一部分。而要研究的最终误差$x-x_0$不过是所有局部因素引起的随机误差之和。这样一来，本例中所关心的误差是大量相互独立的随机变量之和。显然，在实际上人们感兴趣的其他大多数随机变量的情况也是这样的。

由前述的理论研究表明，极其大量的相互独立的随机变

量之和也是一个随机变量，不管各个相加的随机变量的特性如何，只要其中每一个数与总和相比只是一个很小的数，则和的随机变量的分布规律应该接近于某种完全确定模式的规律㊀。

这种确定模式就是正态分布规律。因此可以假定，实际遇到的绝大部分随机变量（大量的相互独立的随机误差相加的总误差）接近于正态分布规律。下面，来认识一下正态分布的一些基本特点。

§ 30　分布曲线的概念

在§15中用图示法描述过分布规律。这种方法很方便，无须查找数据表一眼就可看清要研究的分布规律的一些重要特点。图示法是这样的：在水平直线上分布着某个随机变量的各种可能数值，从某一原点O开始，正数在右，负数在左（图11）。对应于随机变量的每一可能数值，用向上的垂线表示该数值的概率。选定两个方向的比例时应使整个图形方便易读。略看一眼图11就明白，具有最大概率的随机变量的对应值为x_5（负值），并且，随着该随机变量的其他可能数值离开x_5值，其对应的概率也不断地（相当快地）下降。任一区间（α、β）包含的随机变量的概率按加法定理等于该区间所有可能数的对应的概率之和，在几何上，用分布在该区间的各垂线长度之和表示。在图11上，$P(\alpha<x<\beta)=p_1+p_2+p_4+p_5$。如果随机变量的可能数值很大（这种情

㊀参见结论部分。

况在实践中经常碰到），为使图形在水平方向不致延伸过长，在水平方向采用了较大的比例，因此，随机变量可能数值的分布非常密集（图12）。我们看到，所有垂直线的顶端汇成一条曲线，这条曲线称为该随机变量的**分布曲线**。于是，$\alpha<x<\beta$不等式对应的概率用图形表示为在区间（α，β）分布的各垂直线段长度之和。

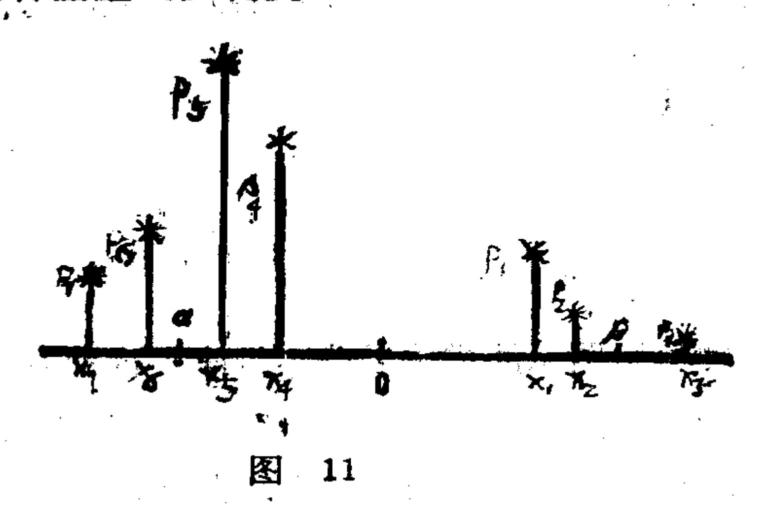


图 11

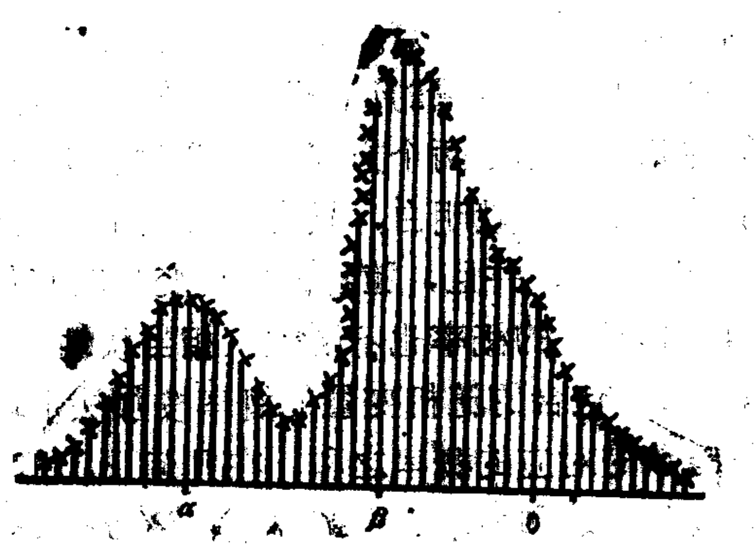


图 12

现在假定随机变量的两个相邻的数值之差始终等于1。这个假定是可能的。如果选择的数值为一列连续正整数，在

选定足够小的比例单位时，上述假定实际上总是能够做到的。这时，每一垂直线段的长度在数值上就等于一个矩形的面积，这个矩形的高为该垂直线段的长，而宽为两相邻线段之间的距离，即宽等于1（图13）。这样一来，不等式$\alpha < x < \beta$对应的概率在几何上可以描绘成如图所示的分布在该区间的矩形面积之和。

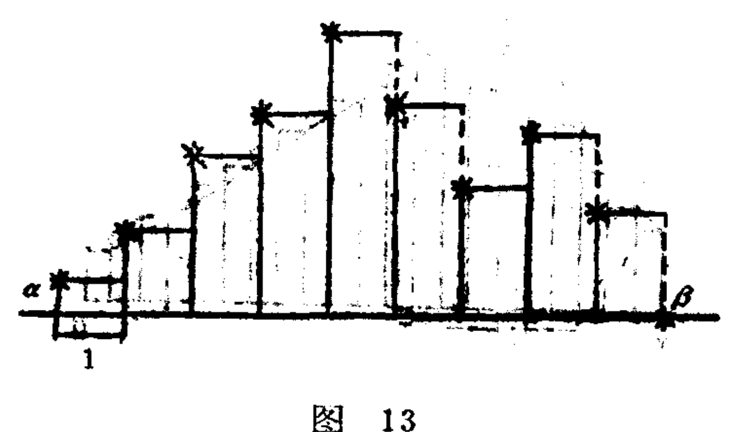


图 13

但是，如果随机变量的可能数值分布很密，如图12所示，则这些矩形面积之和与由α β线段，分布曲线、过α、β点的两条垂线所围成的曲线图形的面积实际上并无差别（见图14）㊀。于是在图14这类曲线图上，该随机变量在任一区间的概率可以简便地用该区间上、分布曲线下的面积来表示。如果该随机变量的分布规律呈这样的曲线图形，那么图上完全不必画出遍布全图的垂直线段，并且单个数值的概率问题本身在此也失去了现实意义。如果可能的值很多（须知，这正是一切曲线图形的基础），那么单个数值的概率照例甚微（实际上等于零），因此完全不会令人关注。例如，在测量

㊀ 当然，这时两个相邻的可能数值之间的距离仍取1。

各相邻两点的距离时，知道测量结果与真实值相差 437cm 的概率是完全无关紧要的。相反，令人关心的问题是偏差在 3～5 m的概率。在其他类似的情况里也是如此，如果随机变量取很多数值，那么，对我们来说，重要的问题不是要知道这些单独数值的概率，而是要知道这些数值中一些完整区间的概率。不过正是这些曲线图形的面积形象而直观地确定了这些概率。

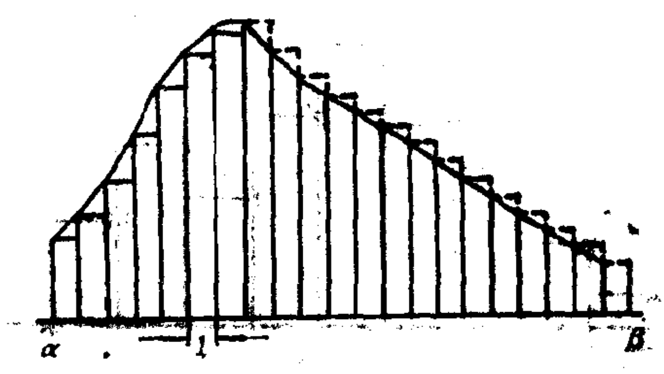


图 14

§ 31 正态分布曲线的特性

鉴于按正态规律分布的随机变量总有无数多个可能数值集合，因此各种正态分布规律都适于用曲线图形来描绘。

图15示出了一些正态分布曲线。尽管这些曲线外表都不相同，但是，可以明显地看出它们有许多共同特点：

1．所有曲线都有一个极值点，当从这一点向左或向右移动时，曲线连续地下降。显然这意味着，当随机变量值偏离其最大概率对应值时，它们对应的概率连续地减少。

2．所有曲线都是轴对称的。对称轴是过极值点的垂

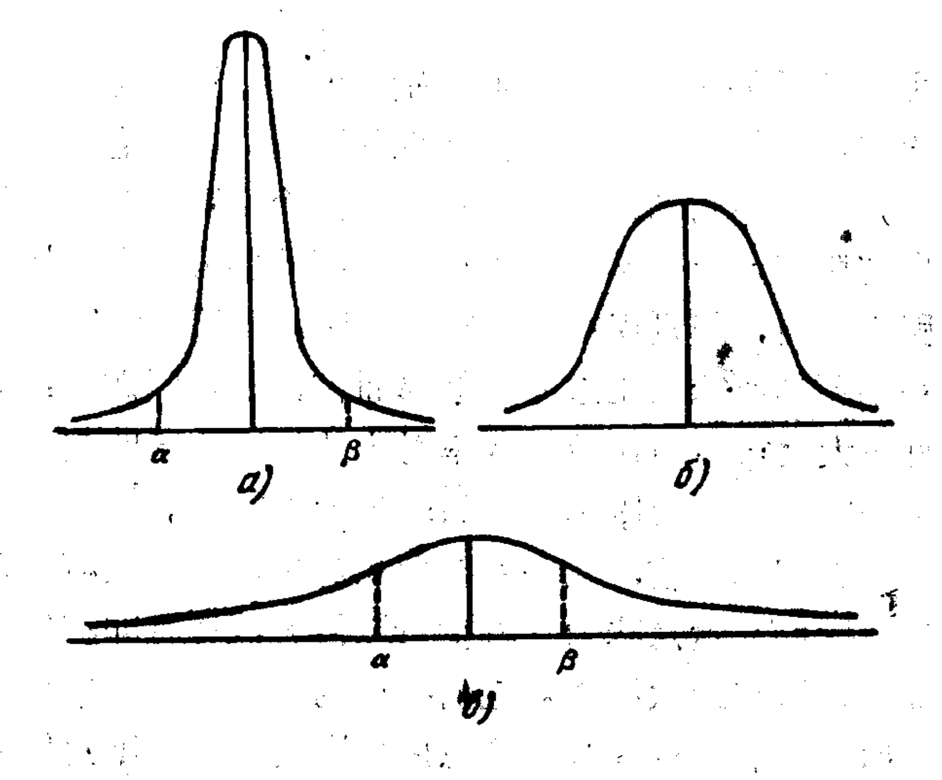


图 15

线。显然这意味着，与最大概率对应值等距离的两个值对应着相同的概率。

8. 所有曲线都呈钟形。在极值点附近，曲线向上凸起，在距极值点某一距离时曲线拐向下凸起，而这个距离（就象曲线最大高度一样）因不同曲线而异㊀。

㊀ 对于懂得高等数学的读者而言，这里指出：描述正态分布规律的曲线方程式为

$$y=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$$

式中，$\exp x=e^x$，$e=2.71828\cdots$，为自然对数的底数，$\pi=3.14159\cdots$，是周长与其直径之比，而a与σ^2是随机变量的均值和方差。读者懂得正态分布规律的解析式，就极易掌握书中的知识。但是，以后的所有内容对不懂高等数学的读者也是易于明了的。

那么不同的正态曲线如何区别呢？要明确回答这个问题，首先应该记住，对所有分布曲线而言，它所包络的面积等于1，因为不管随机变量取何值，这个面积总是等于该随机变量的概率，即必然事件的概率。因此各分布曲线围成的总面积都相同，差别仅在于，在不同区间面积分布情况不同。对于图15所示的正态分布规律而言，根本问题是：总面积中有多大比例集中在最大概率对应值附近的区间内，有多大比例分布在离最大概率对应值较远的区间内。对于图15a）所示的分布规律而言，几乎所有面积都集中在最大概率对应值附近的区间。这意味着随机变量以极大概率（即在绝大多数情况下）取与最大概率对应值接近的值。由于正态分布规律的对称性，最大概率对应值总是与均值重合，则我们可以说：属于a）种分布规律的随机变量的偏离较小，其方差与均方差也小。

相反，在图15б所示的情况里，与最大概率对应值附近区间对应的面积占总面积的较小比例〔如果把图15a与б同样度的区间（α、β）上对应的面积相比较，立刻就看到了差别〕。因此，随机变量取明显偏离其最大概率对应值的概率很大。随机变量的偏离较大，其方差与均方差也大。

显然，情况б介于情况a与б之间。

要想较快地熟悉所有正态分布规律并学会使用它们，较合适的办法是从下列两个基本特性着眼，这两个特性将被详细地介绍而不予以证明，因为要证明首先就要准确地计算正态分布规律，这就要求读者具备高等数学知识。

特性一　如果随机变量x按正态规律分布，则

1）在$c>0$和d为任一常量下，变$cx+d$也按某种

正态规律分布；

2）相反，对于任一正态规律都能找到这样的（唯一的）一对数$c>0$和d，使得变量$cx+d$也按这个正态规律分布。

因此，如果随机变量x按正态规律分布，那么在$c>0$和d的全部可能常数数值下，变量$cx+d$所遵循的分布规律全是正态分布规律。

特性二　如果随机变量x和y相互独立并按正态规律分布，则x与y之和$z=x+y$也按某一正态规律分布。

采用这两个基本特性，就可以较严密地论证正态规律在实践中的一系列特别重要的性质：

1、对于任何两个数a和$q>0$，存在着一个均值为a、均方差为q的唯一的正态分布规律。

事实上，假设x是一个按均值为$\overline{x}$、均方差为Q_x的正态规律分布的随机变量，那么存在着唯一的一对数$c>0$和d，满足随机变量$cx+d$的均值为a、均方差为q的要求，这一论点采用特性一即可证明。如果随机变量x如下表所示：

x_1	x_2	…	x_n
p_1	p_2	…	p_n

则随机变量$cx+d$（式中$c>0$、d暂为任一常数）将对应于下表：

cx_1+d	cx_2+d	…	cx_n+d
p_1	p_2	…	p_n

显然，$\sum\limits_{k} x_k p_k = \overline{x}$

$$\sum_{k} (x_k - \overline{x})^2 p_k = Q_x^2 \text{ ㊀}$$

论题的要求归结为以下两个条件：

$$\sum_{k} (cx + d) p_k = a$$

$$\sum_{k} (cx_k + d - a)^2 p_k = q^2$$

由第一个条件有

$$c \sum_{k} x_k p_k + d \sum_{k} p_k = a$$

或 $\quad c\overline{x} + d = a \qquad\qquad (1)$

由第二个条件有

$$\sum_{k} (cx_k + d - c\overline{x} - d)^2 p_k$$

$$= c^2 \sum_{k} (x_k - \overline{x})^2 p_k = c^2 Q_x^2 = q^2$$

由此（考虑到c＞0）

$$c = q/Q_x \qquad (2)$$

由（1）式有

$$d = a - c\overline{x} = a - \frac{q\overline{x}}{Q_x} \qquad (3)$$

㊀ 符号 $\sum\limits_{k}$ 是 $\sum\limits_{k=1}^{n}$ 的简写。

于是根据给定的数值a与q，利用公式（2）和（3），总是可以且唯一地求出一对数c和d，使得随机变量cx+d遵循均值为a、均方差为q的正态分布规律。论题由此得证。

如果研究范围不限于正态分布，而是所有可能的分布规律，鉴于有非常多的（并且各不相同的）分布规律都具有同样的均值和同样的方差，因此，给出随机变量的均值、方差或均方差，这对于了解它的分布规律只不过是很少的信息。在一般情况下，给出某一随机变量的均值及方差只能颇为近似地描述它的分布规律。

如果只限于研究一些正态分布的话，情况就不同了。一方面，正如刚刚证明过的，关于某一随机变量的均值、方差的任何假设都同该变量遵循正态分布规律的要求相容。另一方面，这也是最主要的方面，如果有根据预先假定某随机变量呈正态分布，那么，给出随机变量的均值和方差，它的分布规律就准确地确定了，这样，随机变量的特性也就完全可知了。知道了这个随机变量的均值和方差，还可以计算出任意区间对应的概率。

2．所有正态分布的中位差（概差）与均方差之比相同。

设有两种任意的正态分布，并设x为遵循第一种正态分布的随机变量。根据基本特性一，存在两个常数c>0和d，使得随机变量cx+d遵循第二种正态分布。用Q_x和E_x分别表示第一个随机变量的均方差和中位差（概差），并用q和e分别表示第二个随机变量的这两种数值。根据概差的定义：

$$P\{|(cx+d)-(c\overline{x}+d)|<e\}=\frac{1}{2}$$

或

$$P\{c|x-\overline{x}|<e\}=\frac{1}{2}$$

或

$$P(|x-\overline{x}|<\frac{e}{c})=\frac{1}{2}$$

根据概差的定义，$\frac{e}{c}$是随机变量x的概差，即$\frac{e}{c}=E_x$，于是$\frac{e}{E_x}=c$。根据公式（2）有$\frac{e}{E_x}=\frac{q}{Q_x}$，于是$\frac{e}{q}=\frac{E_x}{Q_x}$，即两个正态分布的概差与均方差之比相同。

因为根据假设，它们是两种任意的正态分布，于是论题证毕。

这样一来，$\frac{e}{q}$之比是绝对常数，用λ表示。经计算，$\lambda=\sqrt{\frac{2}{\pi}}\approx0.674$。即是说，对于任一正态分布而言，$e=\sqrt{\frac{2}{\pi}}q$。

对于按正态规律分布的随机变量，鉴于e与q两数之间有极其简单的关系式，实际上采用两种偏离程度特征值中的哪一种都一样。上述情况表明（甚至不限于呈正态分布的随机变量），均方差具有其他特征值所没有的一系列简易特性，因此，不论是理论家还是实践家在大多数情况下都选择均方

差作为随机变量偏离程度的度量。前面已指出，炮兵专家几乎总是使用中位差。现在顺便提一下，为什么这种传统方法没有引起任何失误呢？鉴于炮兵科学与实践碰到的随机变量几乎总是按正态规律分布的，因此根据上述比例关系，对于这些随机变量而言，两个特征值中任意用哪一个都一样。

3. 设x和y是相互独立的遵循正态分布规律的随机变量，且$z=x+y$，则

$$E_z=\sqrt{E_x^2+E_y^2}$$

式中E_x、E_y、E_z分别表示变量x、y、z的概差。

不管随机变量x和y的分布规律如何，与§25类似的均方差公式必定存在。当x、y为正态分布时，根据基本特性2，随机变量z也呈正态分布。因此，根据特性2

$$E_x=\lambda Q_x,\ E_y=\lambda Q_y,\ E_z=\lambda Q_z$$

则有

$$E_z=\lambda\sqrt{Q_x^2+Q_y^2}=\sqrt{(\lambda Q_x)^2+(\lambda Q_y)^2}$$

$$=\sqrt{E_x^2+E_y^2}$$

由此可见，在正态分布情况下概差（中位差）也具有均方差同样具有的一个重要特性。

§32 题 解

假定把均值等于零、方差等于1的正态分布称为**基本正态分布**。如果x是遵循基本正态分布的随机变量，则对任一正

数a，可简写成

$$P\{|x|<a\}=\Phi(a)$$

即$\Phi(a)$是遵循基本正态分布的绝对值不大于a的随机变量x的概率。关于不同数值a对应的$\Phi(a)$函数值的准确数据早已算出，并已列成表格。它是计算概率不可缺少的工具。在讲述概率论的所有书籍里都附有此表。读者在本书末尾也可以找到它。手中有了$\Phi(a)$函数数据表，就可以方便准确地对呈正态分布的任何随机变量作一切计算。现举几个实例。

习题1 随机变量x呈正态分布，其均值为$\overline{x}$，均方差为Q_x。求偏差$x-\overline{x}$绝对值不大于a的概率。

设z是遵循基本正态分布的随机变量。根据第121页基本特性一，可以求出c>o和d两个数，使得变量cz+d的均值为$\overline{x}$，均方差为Q_x，即它和该随机变量x一样也遵循同一正态分布规律。因此

$$P(|x-\overline{x}|<a)=P(|(cz+d)-(c\overline{z}+d)|<a)$$

$$=P(c|z-\overline{z}|<a)$$

又根据第123页公式(2)，及$Q_z=1$（对于基本正态分布，方差等于1），因此$c=\frac{Q_x}{Q_z}=Q_x$。于是

$$p(|x-\overline{x}|<a)=P\left(Q_x|z-\overline{z}|<a\right) \quad (4)$$

$$=P(|z|<\frac{a}{Q_x})=\Phi(\frac{a}{Q_x})$$

由于函数$\Phi(\frac{a}{Q_x})$数值可以直接从表中查得，故所提问题得解。这样一来，对遵循任一正态分布的随机变量利用函数表和公式（4）就可以计算出任意偏差界限值的概率。

例1　在车床上加工某种零件。发现其长度是呈正态分布的随机变量，均值为20cm，方差为0.2cm。求零件长度在19.7～20.3cm之间的概率，即求上下偏差都不超过0.3cm的概率。

根据公式(4)及函数表，有

$$P\{|x-20|<0.3\}=\Phi(\frac{0.3}{0.2})=\Phi(1.5)=0.866$$

于是，在该加工条件下制造出的全部零件中，有近87%的零件长度在19.7cm～20.3cm之间，其余13%零件的长度与均值的偏差较大。

例2　在例1条件下，要保证概率在0.95以上，对应的零件长度的精度应是多少？

显然，问题是求满足

$$P\{|x-20|<a\}>0.95$$

的正数a。

例1表明，此处$a=0.3$太小，因为$a=0.3$时，上例不等式的左边小于0.87。根据方程(4)

$$P\{|x-20|<a\}=\Phi(\frac{a}{0.2})=\Phi(5a)$$

首先在函数表中查出$\Phi(5a)>0.95$时的5a值。查表得$5a>1.97$。由此$a>0.394$。

这就是说，长度偏差不超过0.4cm的概率肯定在0.95以上。

例3　人们认为，在一些实际问题中，呈正态分布的随机变量x的偏差不大于其均方差Q_x的三倍。试问这个论断有什么根据呢？

由公式(4)及函数表

$$P\{|x-\overline{x}|<3Q_x\}=\Phi(3)>0.997$$

因此

$$P\{|x-\overline{x}|>3Q_x\}<0.003$$

实际上这意味着，随机变量偏差绝对值大于$3Q_x$的概率平均小于千分之三。这种概率是忽略不计还是必须予以考虑，这取决于问题的内容而不可一概而论。

关系式$P\{|x-\overline{x}|<3Q_x\}=\Phi(3)$显然是下式的局部情况

$$P\{|x-\overline{x}|<aQ_x\}=\Phi(a) \qquad (5)$$

上式由公式(4)导出，适用于呈正态分布的所有随机变量x。

例4 某种零件的平均重量为8.4kg重，已知偏差绝对值大于50g重的概率在每100个零件中有3次。假设零件重量呈正态分布，求重量的概差。

已知

$$P(|x-8.4|>0.05)=0.03$$

式中x为随机选择的零件的重量。由此

$$0.97=P(|x-8.4|<0.05)=\Phi\left(\frac{0.05}{Q_x}\right)$$

查表得知，当$a\approx2.12$时，$\Phi(a)=0.97$

因此

$$\frac{0.05}{Q_x}\approx2.12$$

于是

$$Q_x \approx \frac{0.05}{2.12}$$

根据第126页的公式得到概差为

$$E_x = 0.674Q_x \approx 0.0155\text{kg重} = 15.5\text{g重}$$

例5 炮击时引起炮弹偏离目标有以下三个相互独立的原因：1）确定目标位置的误差；2）瞄准误差；3）更换射手引起的误差（设所有炮弹及气象条件相同）。假设三个误差都呈正态分布，均值为0，它们的概差分别等于24m、8m和12m，求离目标的总偏差不大于40m的概率。

因为根据第126页性质3，随机变量x的总误差的概差为

$$\sqrt{24^2 + 8^2 + 12^2} = 28\text{m}$$

那么总误差的均方差为

$$\frac{28}{0.674} \approx 41.5$$

即 $P(|x| < 40) = \Phi(\frac{40}{41.5}) \approx \Phi(0.964) = 0.665$

因此，出现射击偏差不超过40m的情况约占全部射击中的2/3。

习题2 随机变量x呈正态分布，其均值为$\bar{x}$，均方差为Q_x。求偏差$x - \bar{x}$的绝对值在数a和b（$0 < a < b$）之间的概率。

因为根据加法定理

$$P(|x - \bar{x}| < b)$$

$$= P(|\bar{x} - x| < a) + P(a < |\bar{x} - x| < b)$$

则

$$P(a<|x-\bar{x}|<b)=P(|x-\bar{x}|<b)$$
$$-P(|\bar{x}-x|<a)=\Phi\left(\frac{b}{Q_x}\right)-\Phi\left(\frac{a}{Q_x}\right) \quad (6)$$

所提问题由此得解。

对于绝大多数实际问题来说，经常使用的随机变量$\Phi(a)$函数表是一种烦琐的计算工具。通常只需计算偏差$x-\bar{x}$落在某一区间内的概率即可。因此，对于整个实际问题而言，除了有数据“全表”外，还期望有一种采用公式（6）由“全表”改成的简表。

现举例建立一种比本书末尾的数据表要粗略得多的，但是在许多情况下完全够用的数据表。把随机变量$|x-\bar{x}|$的全部变化范围分为五个部分：1）从0到$0.32Q_x$；2）从$0.32Q_x$到$0.69Q_x$；3）从$0.69Q_x$到$1.15Q_x$；4）从$1.15Q_x$到$2.58Q_x$；5）$2.58Q_x$以上。

利用公式（4）求得

$$P(|x-\bar{x}|<0.32Q_x)=\Phi(0.32)\approx 0.25$$

$$P(0.32Q_x<|x-\bar{x}|<0.69Q_x)$$
$$=\Phi(0.69)-\Phi(0.32)\approx 0.25$$

$$P(0.69Q_x<|x-\bar{x}|<1.15Q_x)$$
$$=\Phi(1.15)-\Phi(0.69)\approx 0.25$$

$$P(1.15Q_x<|x-\bar{x}|<2.58Q_x)$$
$$=\Phi(2.58)-\Phi(1.15)\approx 0.24$$

$$P(|x-\overline{x}|>2.58Q_x = 1-\Phi(2.58)\approx 0.01$$

计算结果可极方便地用图16表示：

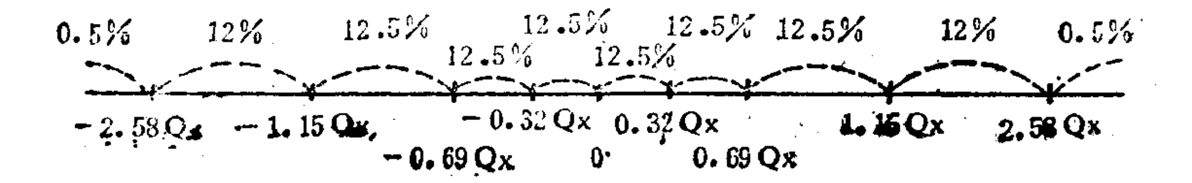


图 16

这里将整个无限长的直线分为十段，五段正的和五段负的。在每一段上都表示出实际上平均有多大百分比的偏差落在该区间上。例如，根据上面的计算，在$(-1.15Q_x, -0.69Q_x)$线段和$(0.69Q_x, 1.15Q_x)$线段共有的偏差应占全部偏差的25%。由于正态分布的对称性，落在这两段的偏差比例常常接近一样，因此两段各约占偏差总数的12.5%。手头上有了这种示意图或与其类似的示意图，对于遵循正态分布的、其均值和均方差为任意值的随机变量，就可以直接找出其偏差分布的基本特点。

最后研究一下如何计算呈正态分布的随机变量落在任一线段上的概率。

习题3 已知随机变量x呈正态分布（平均值为$\overline{x}$，均方差为Q_x），用表计算不等式$a<x<b$的概率，式中a和b（$a<b$）为已知的任意数。

根据数a和b相对于$\overline{x}$的位置，必须研究三种情况：

第一种情况，$\overline{x}\leqslant a\leqslant b$

$$P(\overline{x}<x<b) = P(\overline{x}<x<a)+P(a<x<b)$$

由此

$$P(a<x<b)=P(\overline{x}<x<b)-P(\overline{x}<x<a)$$

$$=P(0<x-\overline{x}<b-\overline{x})-P(0<x-\overline{x}<a-\overline{x})$$

但是，当α>0的任一数时，由于正态分布的对称性

$$P(0<x-\overline{x}<\alpha)=P(-\alpha<x-\overline{x}<0)$$

$$=\frac{1}{2}P(-\alpha<x-\overline{x}<\alpha)=\frac{1}{2}P(|x-\overline{x}|<\alpha)$$

$$=\frac{1}{2}\Phi(\frac{\alpha}{Q_x}) \qquad (7)$$

因此

$$P(a<x<b)=\frac{1}{2}\left\{\Phi(\frac{b-\overline{x}}{Q_x})-\Phi(\frac{a-\overline{x}}{Q_x})\right\}$$

第二种情况，$a\leqslant\overline{x}\leqslant b$

根据公式（7）有

$$P(a<x<b)=P(a<x<\overline{x})+P(\overline{x}<x<b)$$

$$=P(a-\overline{x}<x-\overline{x}<0)+P(0<x-\overline{x}<(b<\overline{x})$$

$$=\frac{1}{2}\left\{\Phi(\frac{\overline{x}-a}{Q_x})+\Phi(\frac{b-\overline{x}}{Q_x})\right\}$$

第三种情况，$a\leqslant b\leqslant\overline{x}$

$$P(a<x<\overline{x})=P(a<x<b)+P(b<x<\overline{x})$$

由此

$$P(a<x<b)=P(a<x<\overline{x})-P(b<x<\overline{x})$$

$$=P(a-\overline{x}<x-\overline{x}<0)-P\ (b-\overline{x}<x-\overline{x}<0)$$

$$=\frac{1}{2}\left\{\Phi\ (\frac{\overline{x}-a}{Q_x})\ -\Phi\ (\frac{\overline{x}-b}{Q_x})\ \right\}$$

三种情况现已求解完毕。我们发现，对于呈任一正态分布的随机变量，采用数据表就可以求出该随机变量落在任一线段上的概率，并可以用这种方法详尽描述该随机变量的分布规律。现举一例。

例 从O点沿直线OX进行射击。炮弹飞行的平均距离等于1200m。假定飞行距离H呈正态分布，其均方差为40 m，求射出的炮弹飞越目标60～80m的百分数。

若炮弹如此飞越目标，则应有$1260<H<1280$。采用习题 3 第一种情况的公式，求得

$$P\ (1260<H<1280)$$

$$=\frac{1}{2}\left\{\Phi\ (\frac{1280-1200}{40})\ -\Phi\ (\frac{1260-1200}{40})\ \right\}$$

$$=\frac{1}{2}\left\{\Phi(2)-\Phi(1.5)\right\}$$

查表得出

$$\Phi(2)\approx0.955,\qquad \Phi(1.5)\approx0.866$$

于是

$$P(1260<H<1280)\approx0.044$$

由此可知，射击的炮弹落在指定范围内的数量略大于4%。

第三部分　随 机 过 程

第十三章 随机过程概论

§33 随机过程的提出

研究自然现象，研究技术、经济或运输系统的过程时，经常遇到用随时间而变化的随机变量来描述这些现象或过程的情况。现举几个例子说明。

众所周知，扩散现象就是一种物质的分子渗透到另一种物质中，各种物质的分子发生混合。现观察某一分子的运动，设在开始时刻$t_0=0$时，被观察的分子位于$(x_0、y_0、z_0)$处，其沿坐标轴的速度分量是$(v_{0x}、v_{0y}、v_{0z})$。在某一随机时刻，该分子与其他分子碰撞，不仅改变了自己的位置，也改变了运动速度和方向。由于我们不可能预先知道分子碰撞的时刻、在任一时间间隔里碰撞的分子数以及与该分子碰撞的其他分子的速度，因此，不能准确地预测该分子的变化。

结果，分子在t时刻的位置由时间的随机函数$x(t)$、$y(t)$、$z(t)$确定。它的速度分量$v_x(t)$、$v_y(t)$、$v_z(t)$也是随时间而变化的随机变量。

现在研究一下由大量元件（电容器、电阻器、二极管、机械零件等等）组成的复杂技术设备。由于种种原因，每个

元件可能丧失本身的工作特性，并处于不能完成其功能的状态。元件的这种状态叫做故障。长期对各种技术设备的观察结果表明，无故障工作时间，即从开始工作时刻到出故障时刻所经过的时间是一个随机变量，因此不可能预先准确指出。现在假定，当某个元件出故障时，有故障的元件立即被同类的新元件替换，并且设备继续工作。试问，在从0到t的时间间隔内必须替换多少个元件？这个数字用n(t)表示，它取决于时间t，是个随机变量。这就是随时间不断改变的随机变量的实例。该随机变量具有这样一个特性：在随机时刻内元件总数不可能减少或改变（即在替换时不减少和改变有故障的元件数）。这种随机函数在可靠性理论——一个广泛采用概率论方法的新的重要的工程科学——中有很大的意义。

在现代工程实践中给工厂、企业供电有很大的学问。某个企业或车间在一定期限内需用多少电能呢？每一给定时刻所需要的功率可能有多大呢？为了使电缆不因输送功率小而浪费，也不因输送功率大于正常工作时所需要的功率而烧毁，那么企业用的电缆应当怎样计算呢？另一方面，由于电缆截面过大耗费金属过多，因而耗费资金也太大，为了不至于此，那么应该怎样计算所需功率呢？当然，要回答这个问题就必须仔细地研究各个机床、机械装置、各种仪器设备乃至连接在同一电缆（电力学的术语称为馈线）上的所有用电户对电能的实际需求情况。在冶金、金属加工、石油开采、化工及其他许多企业里进行过这类研究。现在介绍一下金属加工企业的典型情况。不过，总的情况对其他各种企业也同样适用。

图17表示出车床所需功率 p(t) 的情况。车床工作周期

与不带有效负载的空转周期不断地交替。在特征不同的时间间隔内所需的功率实质上是各不相同的。车床所需功率在空转时接近于零，在工作时急剧增加。由于被加工材料的局部不均匀性，在加工过程中要改变切削速度和切削力，因此急剧增加量也不是常数而有明显波动。同时表明，车床工作周期与空转周期的持续时间的变化完全无规则。深入详尽的研究表明，上述变化是随机性的。于是，又遇到了时间的随机函数。

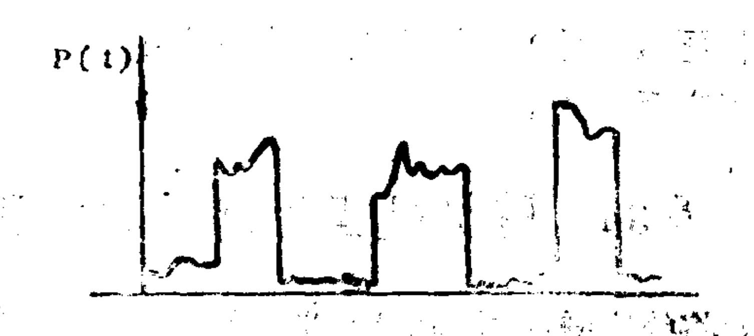


图 17

如果不是记录一台车床而是10～20台一组车床所需的功率，则图17所示的剧烈变化曲线将会变得平缓。所需功率总和不丧失随机性，但具有较大的平稳性。这种现象在很大程度上可用研究大数定理时得知的一些规律来解释。这种新的随机函数的一般形式如图18所示。与平滑过程相联系的事实是：往往在同一时刻，一个车床需要峰值(最大值)功率而另外一些车床需要不大的甚至最小的功率。由前面已知，n个独立随机变量和的方差大约增加$\sqrt{n}$倍。

现在对工厂企业乃至城市电网电力负荷的研究往往都是以上述特性为依据。因此概率论、随机过程理论(即一个独立

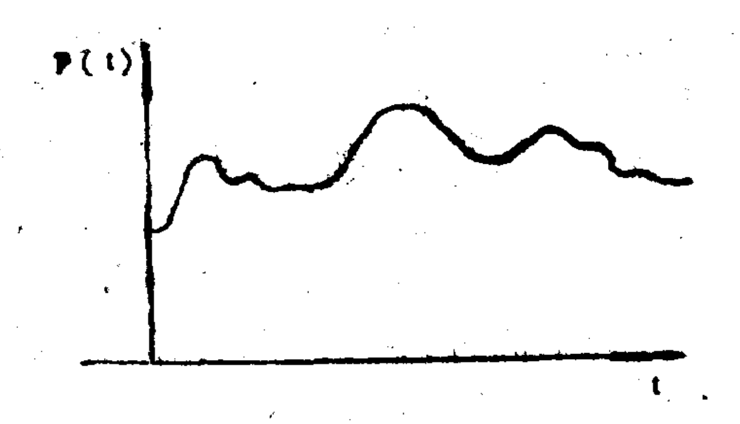


图 18

变量的随机函数理论）的思想方法和数学工具在解决上述问题时得到了广泛应用。

§34 随机过程的概念及类型

下面即将给出随机过程的定义。假定某个随机变量 $\xi(t)$ 与连续变化的参数t有关。尽管这个参数在实际上可能有别的含义，这里仍旧称它为时间。当然，在绝大多数问题里，t实际上是时间。

为了给出随机过程，不仅要学会描述在每一时刻可能采用的一些数值，而且要学会描述以下内容：被采用数值的预期变化，在某一时间内过程可能变化的概率，即将出现的过程演变与其经历时间相互关联程度。不知道这些情况就不能说知道了随机过程。数学上描述随机过程的一般方法是：对于任何正整数n及任何时刻 $t_1, t_2, \cdots, t_n$，可认为下列函数是已知的 $F(t_1, t_2, \cdots, t_n, x_1, x_2, \cdots, x_n) = p\{\xi(t_1) < \xi x_1, \xi(t_2) < x_2, \cdots, \xi(t_n) < x_n\}$

该函数等于在所有选定的时刻 t_i $(i=1, 2, \cdots n)$，不等

式$\xi(t_1)<x_1$同时实现的概率。

随机过程的这种描述方法是通用的，并且原则上可阐述过程随时间变化的一切特征。不过这种方法很粗糙。为了获得较精确的结果，就需要探索另外一条途径，分离出一类重要的随机过程，并为之找出较为完善的适合于计算和建立研究对象数学模型的解析方法。目前已把各种现实的过程划分成几类随机过程，而且对它们的研究工作已经取得了很大的进展。但研究这些问题所必须的数学知识超出了初等数学知识的范围。

在各类典型的随机过程中，**马尔可夫过程**具有特殊意义。该随机过程是为纪念十九世纪末二十世纪初俄国著名数学家A·A·马尔可夫而命名的。马尔可夫首先系统地研究了所谓链式关系特性，而链式关系是建立马尔可夫随机过程概念与理论的原型。

假设过程$\xi(t)$具有下述特性，即对于任意时刻t_0和t，$t_0<t$，从t_0时刻的状态x_0过渡到t时刻的状态x（或过渡到属于某个集合A中的一个状态）的概率只取决于t_0、x_0、t和x（或A），而在t_0时刻以前有关过程的任何情况并不改变这个概率。这些过程以前发展的一切情况正好集中在t_0时刻所达到的x_0状态，且只有通过x_0才影响随后的过程发展。这种过程就称为**马尔可夫过程**。

由于某时刻以前过程进展状态的影响要持续很长时间，因此，乍看起来这种高度概括化的作法（它为马尔可夫过程奠定了基础）与实际要求相差甚远，不过数学本身和数学在生物学、工程技术、物理学及其他知识领域应用中所积累的

经验令人信服地表明，有很多与扩散现象类似的自动化生产管理的现象都可以巧妙地归并到马尔可夫过程公式里。并且，改变状态的概念就可把任一随机过程转变为马尔可夫过程。这就是马尔可夫理论得到广泛发展的一个重要原因。由于马尔可夫过程便于采用较为简单的解析计算方法，因此被广泛地用于许多实际问题的研究之中。

进一步说，研究这种或那种自然现象，研究技术、经济或心理过程时，无论采用任何数学工具事先都必须进行简化，分离出能完全描述其过程的某些特性。现在所关心的不仅是现象的简化，还有它的模型化。建立了现象的数学模型，就有了很多便于对其研究的优越性。首先，它比被研究的现象本身简单，其次，它明确地规定了在实际随机过程（特别是经济和生物现象）中所没有的初始条件和相互关系。如果用相当简单的数学模型研究自然现象，并把所得到的结果与观察该现象的结果进行比较，就可以得出关于该数学模型质量的结论，必要时还可对其进行修正。在建立每个数学模型时都隐约假定：如果每种可能的状态及演变可以用某种选取的数学工具描述时，仅在这时，方可用数学分析法研究某个系统的变化过程。

描述自然现象的最有名的数学模型就是牛顿力学模型。二百五十年来，这种最简单的模型及普通微积分数学计算工具极好地描述了许许多多的随机过程。机械制造业的成就，航天器在近地球轨道的飞行及星际飞行的实现，在很大程度上都是基于广泛采用经典牛顿力学模型。牛顿力学假定质点系统的运动完全可以由其中每个质点的位置和速度来描述。换句话说，已知t时刻的状态就可以单值地计算出该系统在其

他任意时刻的状态。为此，牛顿力学建立了一系列运动方程。

必须注意的是，如果仅仅把t时刻各质点的位置理解为系统的状态，那么这种理解对于单值地确定系统随后的状态是不够的。对牛顿力学来说状态的概念除了位置外还应包括该时刻的速度值。

实际上除了经典力学外所有现代物理中研究的事物的状态都十分复杂。知道系统在某一时刻的状态并不能单值地确定系统的状态本身。马尔可夫过程只是单值地确定在某一时间间隔内从一种状态过渡到另一种状态的概率。如果合适，可把马尔可夫过程看作是对所有经典力学研究过程的广泛综合。

§35 最简单的事件流

在实际遇到的许许多多重要情况里或者在为求知而感兴趣的情况里，必须阐明某类确定事件出现的规律。比如，船舶抵达港口，复杂设备停止工作，过期灯炮的更换、纺纱机断线，放射性物质原子裂变时刻的记录等等事件的规律。理发馆、商店收款处、民航飞机数量、医院必备的病床数量，闸门、路口、桥梁的流通能力等等公共服务事业的计算工作与这类事件流量的研究紧密相关。几年前，人们曾仔细研究了以下情况：飞机飞抵大型机场、货船抵达指定港口、急救医疗站电话呼叫频率、电话局用户呼叫频率等等。观察结果表明，在上述所有情况下这些事件的出现都可以用非常接近真实情况的规律来描述。

用t表示所关心的时刻，并假定$P_k(t)$是流通事件k在该时刻出现的概率。当k＝0，1，2，…时，可以相当准确地建立下列等式

$$P_k(t)=\frac{(\lambda t)^n}{k!}e^{-\lambda t} \quad (1)$$

式中λ为一正常数，表示流通事件出现的“频度”。若在t时刻内没有一个流通事件出现，则其概率等于

$$P_0(t)=e^{-\lambda t} \quad (2)$$

在分子物理学里，需要研究在某时间间隔t里某个分子不与其他分子碰撞的概率。在论述这类课题的物理书籍里表明，这种概率正好等于$e^{-\lambda t}$。如果在本例中把该分子与其他分子碰撞的时刻看作事件流，那么人们恰恰是求在长为t的时间间隔里不发生一次流通事件的概率。

在如此众多性质各异的现象里出现同样的规律性，人们自然联想必定有某种共同的原因。实际表明确实存在一些深奥原因，正由于这些原因，使得在极纷繁的条件下也必定出现上面所说的规律性。本世纪初著名物理学家A·爱因斯坦和M·斯莫努霍夫斯基在研究布朗运动时发现了存在规律性的第一组条件。

假定所研究的事件流具有以下三个特性：

1. **平稳性** **即是在有限数量的互不重叠的时间间隔内出现相应的事件$K_1, k_2, \cdots, k_n$的概率只取决于时间间隔的数量和时间间隔的长短。而且在时间间隔（T，t＋T）内出现K个事件的概率与T无关，只是k和t的函数。**

2. **无后效性** **即在时间间隔（T，T＋t）内，k个事件流出现的概率与在该时刻之前有多少个事件以及它们如**

何出现无关。这意味着被研究的事件流具有马尔可夫特性。

3. **单一性** 即在很短的时间间隔内实际上不可能出现两个或两个以上事件。

满足这三条的事件流叫做最简单的事件流。

不难证明，最简单的事件流完全可由等式（1）求出。

最简单的事件流正可以用于求任意随机时间间隔的瞬时流，而且相邻瞬时的间隔大于t的概率可用公式（2）求出。这种最简单的事件流的定义也时常用于解决很多理论问题和实际问题。

直接检验是否具备三个条件（平稳性、无后效性及单一性）时常难以办到，因此，重要的是寻找另外一种条件，以便依此为基础得出事件流是最简单的或接近于最简单的结论。这种条件在一些研究者的著作中可以找到。其主要内容如下。

假定感兴趣的事件流是数量极多的、彼此独立的平稳事件流之和，而且其中每一个事件流对其总和影响又很小。由于算术方法的局限性保证了总事件流的单纯性，因此，**总事件流近似于最简单的事件流。**

现代概率论创始人之一A·Я·辛钦用一般形式证明了这一理论。该理论为实际应用提供了基本原理。

实际上，这个理论使我们能从事件流的一般结构中得出一些重要的结论。例如，从电话局的呼叫事件流中，可以研究大量独立事件流（其中每一个对总和的影响很小）之总和，这种事件流应接近于最简单的。同样，抵达某一港口的货船流是由从其他各港出发的大量的事件流所组成。因此，货船流也应接近于最简单的事件流。实际情况也的确如此。

其他有现实意义的例子举不胜举。

§36 批式服务的理论问题

本节研究的问题是许多重要实际情况中的典型问题。首先，从工厂、商店、仓库的工作人员，电话网路的设计人员经常面临的实际问题出发，来描述这个问题。

为了满足居民的需要建立了理发馆、电话局、医院、牙科诊所等行业，需要服务的时刻是随机的，服务时间的长短也是随机的，试问，如果设立n个服务点，那么满足顾客需求的情况如何呢？

上述条件极恰当地反映了实际情况。确实不可能指出顾客在何时去理发馆或去牙科诊所。同时也知道，顾客为得到必要的服务经常要排队等候，当然也有时不排队等候。看起来，要完成同样一种（工作）作业实际上需要不同的时间，如在治疗牙齿时医师根据牙齿的症状或者清洗、或者换药，有时在一次诊治中还进行全口拔（镶）牙。

当然，无论是顾客还是企业领导人，首先关心的是一些服务质量指标，即顾客等候服务的平均持续时间及服务设备的负荷程度等等（这里假定已知从预约到服务的平均速度以及服务的平均速度）。

要解决这一问题需从以下假设出发：1）服务需求流是最简单的事件流；2）服务时间的长短是随机的，并且服务所用时间不少于t的概率等于$e^{-\nu t}$，其中$\nu>0$并为一常数；3）每一需求由一个设备来满足，每一设备工作时只满足一个需求；4）如果服务时要排队，那么，空间的设备应立即为依

次等候的用户服务。

用$P_k(t)$表示在t时刻有k个需求得到满足的概率。根据我们规定的条件，当k为任意值，即k＝0，1，2，…时，其对应的概率可以求得。不过，准确的公式很烦琐，因此实际上不用它，而宁愿用由烦琐公式导出的适于确定工作方式的公式。这后一种公式特别简单，其形式为：

当$1\leqslant k\leqslant n$时

$$p_k=\frac{\rho^k}{k!}p_0, \tag{3}$$

当$k\geqslant n$时

$$p_k=\frac{\rho^k}{n!\ n^{n-k}}p_0 \tag{4}$$

式中

$\rho<n$时

$$p_0=\left(\sum_{k=0}^{n}\frac{\rho^k}{k!}+\frac{\rho^{n+1}}{n!\ (n-\rho)}\right)^{-1} \tag{5}$$

$\rho\geqslant n$时

$$P_0=0$$

在上述公式里，假定$\rho=\lambda/\nu$。

注意到当$\rho\geqslant n$时，概率$P_0=0$，由公式（3）、（4）表明，当k为$k\geqslant1$的任一值时

$$d_k=0$$

换句话说，在确定的服务程序下，当$\rho\geqslant n$时，在某个系统里碰到有限个需求的概率为零，即是说在该系统里将会碰到无穷多个需求，并且组成无数个排队顺序，这种概率为1。这就意味着，在$\rho\geqslant n$的所有情况下，要求服务的队伍随时

间的增加而无限地增长。

这个结论有很大的实际意义，这是由于在设计这些必须的设施（机场着陆面积、海港码头、医院病床、企业里工具分发点、商店收款处等等）的数量时，经常采用了错误的前提，从系统的“理想的”生产能力出发，即以服务用设施的数量乘以这些设施的使用寿命再除以平均每次服务的持续时间。但是用户对服务的需求是不平衡的，因此上面的计算方法必然引起排队，浪费时间、消耗设备以至失去主顾。

当然，学会使用批式服务理论方法不仅能避免服务系统负载过量的危害，而且也能避免服务系统中设备数量过剩的损失。现可以举出许多例子来说明批式服务理论已经成为各个社会活动行业（如电话局、企业维修队、大型机场或大流通汽车隧道运输）计算时的必要工具。现在，批式服务理论在设计制造现代化计算机、信息系统中，在核物理和生物学等领域中都具有越来越大的意义。

§37 关于可靠性理论问题

在近四分之一世纪，全世界的注意力都集中在可靠性理论这一新的科学领域。其目的在于研究工业产品在设计、制造、检验、运输、贮藏及使用中应该遵循的一般规则，以确保其最大使用效率。当然，根据元件的可靠性，研究复杂产品及复杂技术系统可靠性的计算方法，也是可靠性理论要研究的内容。

由于人类的全部生活都直接或间接地使用着各类技术设备与系统，因此，该理论的实际重要意义是毋庸置疑的。每

天人们坐汽车或电车上班，使用房间照明开关，使用由远处压水机经管道输送来的自来水，在医院里利用各种设备帮助病人恢复身体机能。又如在作完肾脏手术之后采用特殊的器械——人工肾脏，在肾脏功能的必要恢复时期，人工肾脏代行它的功能；每年飞机运载成千上万的旅客到世界各地。在上述所有情况里我们特别关注的是：使用的技术设备能在整个工作时期绝对正确地完成赋予它的功能。破坏这个要求就要导致不可挽救的后果。的确，当飞机发动机在飞行中出故障时，飞机就可能坠毁；人工肾脏在停止履行自己的功能时，就可能威胁到术后病人的生命；送水站压水机的可靠性不足造成事故，就会破坏市内供水。结果，市民就不能制作食品，无水饮用和洗漱；手术被迫停止，正常的医疗受到破坏；街上因无水洒扫而尘土飞扬，如此不一而足。

所有这些问题似乎与概率论没有任何关系，而应由工程师、设计师、工厂工人与使用者解决。然而并非如此，在选择合理的产品质量试验计划、从试验结果中得出结论、在设备预防性的监测、维修的最佳周期的定量计算方面，数学家都要担负相当大的责任。由此可见，产品的所有基本工作特性都具有概率特征。例如，在同一工厂用同一原料在同样条件下生产出来的同一种产品，其无故障工作寿命却千差万别。如果回忆一下电灯泡从启用到扔进垃圾箱的全部寿命是何等不同这一事实（有时一个电灯泡工作好几年，而有时虽只用了几天因灯丝烧断也只好更换），就足以得出上述结论。

长期的观察及无数专门试验令人信服地表明，虽然不可能精确地指出某个零件使用寿命有多长，但能评估该零件使

用时间不小于给定数值t的概率。

这样一来，概率论确实渗透到可靠性理论的所有问题之中，并且是解决可靠性理论问题的基本方法之一。

现在来研究一个不太复杂的概略计算的问题。为不使问题复杂化，简明扼要地阐述一下所碰到的实际问题。

众所周知，在自然界没有绝对可靠的元件和产品。每个元件无论其性能多么完善，也会随时间而改变。因此要提高产品的可靠性必须采用以下方法：降低使用条件，寻找较为完善的材料、新的设计构造或连接方式。提高可靠性的一个最普遍的方法是备份法。备份法的实质是在系统内置入过剩的元件，即零件甚至整机，当主要元件丧失工作能力时它们取而代之。比如，为了保障铁路列车畅通，在枢纽站保留有备用的蒸汽机车与电气机车。在大型电站里必须有备用发电机。在特别重要的供电线路上采用并行线路（它们在正常状况下负荷不饱满）。每个汽车除有四个轮胎外，还有一个备用轮胎。

假定在t时间内应有n个设备同时工作。在这个期间，其中任一设备无故障工作的概率等于p（所有设备都为同一概率，且各设备彼此无关）。如果一个设备发生故障，就会导致全系统发生故障（例如，一个汽车轮胎瘪气就引起整个汽车出故障）。根据伯努利定理，系统无故障工作的概率等于p^n。如果在系统内除了有n个基本设备外还有m个处在负载状态（即与基本设备处于同一状态）的备用设备。当系统处于其工作设备数小于n个状态时，认为它发生了故障。根据概率加法定理，所求概率等于

$$\sum_{i=0}^{m} C_{m+n}^{n+i} P^{n+i}(1-p)^{m-i}$$

现举一不太复杂的数字说明上式。设$n=4$，$m=1$，$p=0.9$。不难算出，无备件系统无故障工作概率等于0.6561。如果有一个备份件，则概率就等于0.9185。这样，四个元件的系统有了一个备份件，无故障工作概率就几乎提高了1/2倍。如果具有两个备份件，系统无故障工作的概率就会提高到0.9841。这就是为什么在电站有一个贮备发电机就几乎可完全保证电站不间断工作的缘故。

如果采用所谓**"功能复原贮备件"**，系统可靠性就会提高许多倍。哪个设备有故障就立即送去修理，当功能恢复后又重新作为备份件。这种方法能多倍地提高贮备件系统的可靠性。

上面只是研究了简化的备份件理论问题。研究现实条件下备份件理论需要掌握较为复杂的数学知识，首先是随机过程理论。现在，可靠性理论的一些重要问题已经得到解决，但是尚有许多问题远未最终解决，即使解决了的问题也远不令人满意。在某些简化的条件下进行系统性的研究可以逐渐解决这些问题，也可以在较为现实的条件下开阔研究这些问题的新途径。

结 束 语

近几十年来，概率论已成为发展最快的数学分支之一。它的新的理论研究成果为在自然科学和实际生活中采用概率论方法开拓了新的前景。同时，对自然现象乃至生产、技术、经济及其他过程深入细致地研究也推动了概率论在探索新方法、新规律方面的进一步发展。概率论是一门与生活、与其他科学密切相关的数学分支，它与自然科学和技术的发展并行不悖。读者不应将上述事实理解为，目前概率论只是解决实际问题的补助手段或工具。这是绝对不可以的。事实上，在近五十年内概率论已经发展成为一门有其研究对象、研究方法的严谨的数学科学了。而且概率论也和其他数学一样，其最重要的、理所当然的任务是解决自然界的紧迫问题。

概率论产生于十七世纪中叶。在费尔马（1601～1665）、帕斯卡（1623～1662）和格尤更斯（Гюйгенс）（1625～1695）等学者的著作中已经出现了随机事件概率、随机变量的数学期望等概念的萌芽。当时他们研究的目的是解决与赌博有关的问题。不过，他们已经清醒地认识到了这些概念对于研究自然现象的重要性。例如格尤更斯在《论赌博计算》一文中写道："读者发现，它不仅与游戏有关，而且为很有趣、很深的理论奠定了基础"。这之后，对概率论发展有重大影响的学者之中，值得一提的是雅可夫·伯努利（1654～1705）——本书中已经出现过他的名字，还有A·姆阿弗尔（1667～1754），T·巴叶斯（1763年逝世），П·拉普拉斯

(1749～1827)，K·高斯(1777～1855)及C·泊松(1781～1840)。

概率论的巨大发展与俄国科学家的创造性成就紧密相连。在上一世纪，当西方概率论研究走进死胡同时，俄国天才的数学家П·Л·车比雪夫开阔了新的途径，较全面地研究了独立随机变量的序列。车比雪夫本人及其学生A·A·马尔可夫，A·M·李雅普诺夫采用这种方法获得了奠基性的巨大成果(大数定理，李雅普诺夫理论)。

读者已经知道了大数定理，因此现在主要介绍一下概率论中另一个最重要的假设——李雅普诺夫理论(有时称为概率论的中心极限定理)。

研究表明，大量现象出现都是随机性的，其影响因素相当多。被研究的现象常常受到大量随机因素的独立影响，而每一个因素对整个过程只有非常微小的影响。每个因素的影响用随机变量 ξ_1，ξ_2，…，ξ_n表示，所有因素对该现象影响之和为$s_n = \xi_1 + \xi_2 + \cdots + \xi_n$。因为计算每个因素的影响(换句话说，给出随机变量 ξ_k的分布函数)，甚至简单地描述这些因素的影响实际上都不可能，因此，找出一些与每个因素的特性无关、而又能研究这些因素总影响的方法是多么重要啊！要解决这个问题，用普通的研究方法是无能为力的，需要寻求一些新方法，对于这些新方法来说，某一现象受到的多种影响因素并不妨碍问题求解，相反还使问题易于解决。这类方法应该用大量单个作用的因素来弥补对每一因素的无知。

俄国科学院院士П·Л·车比雪夫(1821～1894)、A·A·马尔可夫(1856～1922)和A·M·李雅普诺夫

（1857～1918）在他们的论著中建立了中心极限定理，该定理表明，如果作用因素 ξ_1， ξ_2，…，ξ_n彼此独立，这些因素的数量n很大，而且每个因素的影响与所有因素的总影响 相比非常微小，那么总和s_n的分布规律就近似呈正态分布。

现举例说明：

当用武器射击目标时，子弹弹着点不可避免地要偏离瞄准点。这就是众所熟知的子弹偏离现象。由于该偏离现象是许多独立因素影响的结果（例如，弹壳、弹头外圆尺寸不精确，弹头的材料密度的变化，各弹头内炸药量的微小变化，瞄准时肉眼不能发现的细小误差，在各种射击情况下大气状况的微小变化等等），其中每项因素对子弹弹道只有很小的影响，那么根据李雅普诺夫定理，偏离应遵循正态分布规律。这种情况在射击理论中都考虑到了，并成为制订射击规范的基础。

当我们对某种物理常数进行观测时，不可避免地会有大量因素影响观测结果，虽然单个因素可以不考虑，但它却影响测量精度。测量仪器的误差就属于这种情况，即不在不同的大气、温度、机械或其他的因素影响下，测量仪器的读数可能有微小的变化。由观察者视觉、听觉特性引起的误差，甚至观察者心理与体质状况的微小变化引起的误差也属于这种情况。因此实际测量误差是大量的、数值很小的、彼此独立的随机误差的总合。由李雅普诺夫定理可以再次认定，观测误差将遵循正态分布规律。

类似的例子可以随意举出一些。例如，与其他分子大量碰撞的气体分子的位置和速度，扩散物质的数量，在大批生产时机械零件与其名义尺寸的偏差，动植物及其器官组织的生

长分布特性等等。

统计物理学及其他技术的发展向概率论提出了许许多多不属于经典数学范畴的新问题。当物理学或工程技术着重研究随时间变化的现象（即随机过程）时，概率论尚没有解决此类问题的普遍适用的或局部适用的方法。因此，研究随机过程的一般性理论、即研究随一个或多个不断变化的参数而变化的随机变量的理论就十分必要了。

苏联数学家A·H·柯尔莫哥洛夫和A·Я·辛钦为随机过程的一般理论奠定了基础。

在本世纪初，A·A·马尔可夫研究了相关随机变量的序列，这个著名的理论称为马尔可夫链。该理论起初只作为一个数学课目，到本世纪二十年代，它已成为物理学家研究自然现象的有效工具。递之后，许多学者对马尔可夫链理论都作出了很大贡献。

在二十年代，A·H·柯尔莫哥洛夫，E·E·斯努茨基、A·Я·辛钦及保尔·列夫找到了概率论与集合论之间的密切联系以及集合论与实变函数论的一般概念。在这之前不久，Э·鲍列里也找到了这一规律。他们的这一发现特别重要。根据这种方法就可以彻底解决车比雪夫提出的经典问题。

最后应该指出，C·H·别尔宁士捷因、A·H·柯尔莫哥洛夫等人在建立逻辑严谨的涉及自然科学、技术及其他各知识领域的概率论方面的贡献。虽然上述作者的著作为建立概率论逻辑基础起了很大的推动作用，但是有关这些问题的研究工作仍在不断地加强。其原因就是人们希望了解自然界本身的随机性，希望在自然现象的随机性与其必然性之间建

立某种关系。这个重大的哲学课题即使在今天不能完全解决，也要鼓励人们向它进军。

苏联、美国、法国、英国、瑞士、日本、匈牙利的科学研究对今日概率论的迅猛发展起了巨大作用。加之全世界在迫切要求解决实际问题的影响下，对这门科学的兴趣也大为增加。

如同每门发展中的知识领域一样，概率论也需要科技界新生力量不断地推动它向前发展。在青年研究人员面前展现出从事创造性研究、施展自己才华的广阔天地。但是要使才能充分发挥出来就不能陷入事务主义的狭窄胡同，而应对概率论的所有方面深感兴趣，即对概率论的逻辑基础问题、与其他数学分支关联的问题，对自然科学（物理学、生物学、化学等）、工程技术、生产、经营管理以及其他理论与实践中提出来的许多新问题都有着强烈的探索求知欲望。

附　　录

Φ(a)　函数数值表

a	Φ(a)	a	Φ(a)	a	Φ(a)
0.00	0.000	0.20	0.159	0.40	0.311
0.01	0.008	0.21	0.166	0.41	0.318
0.02	0.016	0.22	0.174	0.42	0.326
0.03	0.024	0.23	0.182	0.43	0.333
0.04	0.032	0.24	0.190	0.44	0.340
0.05	0.040	0.25	0.197	0.45	0.347
0.06	0.048	0.26	0.205	0.46	0.354
0.07	0.056	0.27	0.213	0.47	0.362
0.08	0.064	0.28	0.221	0.48	0.369
0.09	0.072	0.29	0.228	0.49	0.376
0.10	0.080	0.30	0.236	0.50	0.383
0.11	0.088	0.31	0.243	0.51	0.390
0.12	0.096	0.32	0.251	0.52	0.397
0.13	0.103	0.33	0.259	0.53	0.404
0.14	0.111	0.34	0.266	0.54	0.411
0.15	0.119	0.35	0.274	0.55	0.418
0.16	0.127	0.36	0.281	0.56	0.425
0.17	0.135	0.37	0.289	0.57	0.431
0.18	0.143	0.38	0.296	0.58	0.438
0.19	0.151	0.39	0.303	0.59	0.445

续表 1

a	Φ (a)	a	Φ (a)	a	Φ (a)
0.60	0.451	0.85	0.605	1.10	0.729
0.61	0.458	0.86	0.610	1.11	0.733
0.62	0.465	0.87	0.616	1.12	0.737
0.63	0.471	0.88	0.621	1.13	0.742
0.64	0.478	0.89	0.627	1.14	0.746
0.65	0.484	0.90	0.632	1.15	0.750
0.66	0.491	0.91	0.637	1.16	0.754
0.67	0.497	0.92	0.642	1.17	0.758
0.68	0.504	0.93	0.648	1.18	0.762
0.69	0.510	0.94	0.653	1.19	0.766
0.70	0.516	0.95	0.658	1.20	0.770
0.71	0.522	0.96	0.663	1.21	0.774
0.72	0.528	0.97	0.668	1.22	0.778
0.73	0.535	0.98	0.673	1.23	0.781
0.74	0.541	0.99	0.678	1.24	0.785
0.75	0.547	1.00	0.683	1.25	0.789
0.76	0.553	1.01	0.688	1.26	0.792
0.77	0.559	1.02	0.692	1.27	0.796
0.78	0.566	1.03	0.697	1.28	0.800
0.79	0.570	1.04	0.702	1.29	0.803
0.80	0.576	1.05	0.706	1.30	0.806
0.81	0.582	1.06	0.711	1.31	0.810
0.82	0.588	1.07	0.715	1.32	0.813
0.83	0.593	1.08	0.720	1.33	0.816
0.84	0.599	1.09	0.724	1.34	0.820

续表 2

a	Φ(a)	a	Φ(a)	a	Φ(a)
1.35	0.823	1.60	0.890	1.85	0.936
1.36	0.826	1.61	0.893	1.86	0.937
1.37	0.829	1.62	0.895	1.87	0.939
1.38	0.832	1.63	0.897	1.88	0.940
1.39	0.835	1.64	0.899	1.89	0.941
1.40	0.838	1.65	0.901	1.90	0.943
1.41	0.841	1.66	0.903	1.91	0.944
1.42	0.844	1.67	0.905	1.92	0.945
1.43	0.847	1.68	0.907	1.93	0.946
1.44	0.850	1.69	0.909	1.94	0.948
1.45	0.853	1.70	0.911	1.95	0.949
1.46	0.856	1.71	0.913	1.96	0.950
1.47	0.858	1.72	0.915	1.97	0.951
1.48	0.861	1.73	0.916	1.98	0.952
1.49	0.864	1.74	0.918	1.99	0.953
1.50	0.866	1.75	0.920	2.00	0.955
1.51	0.867	1.76	0.922	2.01	0.956
1.52	0.871	1.77	0.923	2.02	0.957
1.53	0.874	1.78	0.925	2.03	0.958
1.54	0.876	1.79	0.927	2.04	0.959
1.55	0.879	1.80	0.928	2.05	0.960
1.56	0.881	1.81	0.930	2.06	0.961
1.57	0.884	1.82	0.931	2.07	0.962
1.58	0.886	1.83	0.933	2.08	0.962
1.59	0.888	1.84	0.934	2.09	0.963

续表8

a	Φ (a)	a	Φ (a)	a	Φ (a)
2.10	0.964	2.35	0.981	2.60	0.991
2.11	0.965	2.36	0.982	2.61	0.991
2.12	0.966	2.37	0.982	2.62	0.991
2.13	0.967	2.38	0.983	2.63	0.991
2.14	0.968	2.39	0.983	2.64	0.992
2.15	0.968	2.40	0.984	2.65	0.992
2.16	0.969	2.41	0.984	2.66	0.992
2.17	0.970	2.42	0.984	2.67	0.992
2.18	0.971	2.43	0.985	2.68	0.993
2.19	0.971	2.44	0.985	2.69	0.993
2.20	0.972	2.45	0.986	2.70	0.993
2.21	0.973	2.46	0.986	2.72	0.993
2.22	0.974	2.47	0.986	2.74	0.994
2.23	0.974	2.48	0.987	2.76	0.994
2.24	0.975	2.49	0.987	2.78	0.995
2.25	0.976	2.50	0.988	2.80	0.995
2.26	0.976	2.51	0.988	2.82	0.995
2.27	0.977	2.52	0.988	2.84	0.995
2.28	0.977	2.53	0.989	2.86	0.996
2.29	0.978	2.54	0.989	2.88	0.996
2.30	0.979	2.55	0.989	2.90	0.996
2.31	0.979	2.56	0.990	2.92	0.996
2.32	0.980	2.57	0.990	2.94	0.997
2.33	0.980	2.58	0.990	2.96	0.997
2.34	0.981	2.59	0.990	2.98	0.997

续表 4

a	Φ(a)	a	Φ(a)	a	Φ(a)
3.00	0.997	3.50	0.9995		
3.10	0.998	3.60	0.9997	4.00	0.99994
3.20	0.999	3.70	0.9998		
3.30	0.999	3.80	0.99986		
3.40	0.999	3.90	0.99990	5.00	0.99999994